

# **ESCUELA POLITÉCNICA NACIONAL**

## **FACULTAD DE CIENCIAS**

### **DETERMINANTES OBSERVABLES Y NO OBSERVABLES DE LOS DIFERENCIALES SALARIALES ENTRE EL SECTOR PÚBLICO Y PRIVADO EN EL ECUADOR PERÍODO 2011: UN ANÁLISIS EMPÍRICO DE DESCOMPOSICIÓN DE LOS INGRESOS**

**PROYECTO DE TITULACIÓN PREVIO A LA OBTENCIÓN DEL TÍTULO DE  
INGENIERA EN CIENCIAS ECONÓMICAS Y FINANCIERAS**

**ERIKA MARIBEL GUERRERO MENA**

eri\_gm@hotmail.es

**DIRECTOR: Ec. DANIEL IVÁN TORRESANO MELO Msc.**

dtorresano@sbs.gob.ec

**CODIRECTOR: Mat. ALFONSO JUAN ELÍAS CASTRO BALAREZO Msc.**

alfonso.castro@epn.edu.ec

**Quito, 2013**

## **DECLARACIÓN**

Yo, Erika Maribel Guerrero Mena, declaro que el trabajo aquí descrito es de mi autoría; que no ha sido previamente presentada para ningún grado o calificación profesional; y, que he consultado las referencias bibliográficas que se incluyen en este documento.

La Escuela Politécnica Nacional puede hacer uso de los derechos correspondientes a este trabajo, según lo establecido por la Ley de Propiedad Intelectual, por su Reglamento y por la normatividad institucional vigente.

---

**Erika Maribel Guerrero Mena**

## **CERTIFICACIÓN**

Nosotros, Daniel Torresano y Alfonso Castro, certificamos que el presente trabajo fue desarrollado por Erika Maribel Guerrero Mena, bajo nuestra supervisión.

---

**Ec. Daniel Torresano Msc.**

**DIRECTOR**

---

**Mat. Alfonso Castro Msc.**

**CODIRECTOR**

## **AGRADECIMIENTOS**

Agradezco primeramente a Dios por brindarme sabiduría, fortaleza y amor; pero sobre todo le doy gracias por cubrirme con su manto de bendiciones cada instante de vida, pues él ha sido mi apoyo incondicional para culminar mi carrera y ayudarme a crecer como persona.

Mi gratitud a mis padres, Amparito y Alonso, ya que sin su apoyo emocional y económico no hubiese podido cumplir este sueño.

Gracias Msc. Daniel Torresano por compartir con paciencia tus valiosos conocimientos, por tu interés, dedicación, guía y apoyo para que esta investigación se lleve a cabo. En especial, te agradezco por tus lecciones de vida y consejos que me enseñaron a ser mejor ser humano. Muchas Gracias Dany.

Gracias Mat. Alfonso Castro por aportar con sus valiosos conocimientos para que mi investigación culmine con éxito.

Agradezco a la Escuela Politécnica Nacional, en especial a sus profesores que han impartido sus conocimientos, los cuales quedaran gravados en mi mente para ser una excelente profesional.

Además, agradezco a todos mis amigos y amigas, pero en especial agradezco a Alex, Andre, Karencita, Adrián y Bolo que estuvieron a mi lado compartiendo buenos y malos momentos a lo largo de toda mi carrera.

Y agradezco a todos ustedes que se interesaron en mi trabajo de investigación. Muchas Gracias.

## DEDICATORIA

“Son felices los que sueñan y están dispuestos a sacrificarse para que sus sueños se hagan realidad”. Autor: Desconocido.

Dedico esta Tesis y especialmente todos estos años de estudio, esfuerzo y dedicación a Dios quien ha sido mi fortaleza para caminar en esta vida. También mi dedicatoria se dirige a mi padre Alonso, mi madre Amparito, mi hermano Cristian y mi abuelita Aurora; pues ellos han conformado mi puntal de amor, apoyo, comprensión, seguridad y confianza para no detenerme en los senderos difíciles que el ser humano debe recorrer.

Para mis padres Alonso y Amparito porque juntos me enseñaron amar y sentir el verdadero amor que es el de los padres hacia su hija. Además, con su ejemplo de vida, sus consejos y paciencia me indicaron el camino del bien. Gracias papitos porque lo que estoy logrando es también gracias a Uds.

Para mi hermano Cristian, porque quiero ser un ejemplo para ti y que veas que quienes nos sacrificamos y damos nuestro mayor esfuerzo logramos cumplir nuestros sueños. Ánimos hermanito en esta nueva etapa de tu vida, tu lograrás lo que te propongas.

Para ti abuelita Aurora, quien has sido mi inspiración de vida porque el amor que me brindaste desde niña siempre estará guardado en lo más profundo de mi corazón.

Les quiere,

Erika Maribel

# ÍNDICE DE CONTENIDO

<b>LISTA DE GRÁFICOS</b> .....	i
<b>LISTA DE TABLAS</b> .....	ii
<b>LISTA DE ANEXOS</b> .....	iii
<b>RESUMEN</b> .....	iv
<b>ABSTRACT</b> .....	v
<b>CAPITULO I</b> .....	1
<b>INTRODUCCIÓN</b> .....	1
1.1. OBJETIVOS DE LA INVESTIGACIÓN .....	4
1.1.1 OBJETIVO GENERAL .....	4
1.1.2 OBJETIVOS ESPECÍFICOS .....	4
1.2. JUSTIFICACIÓN.....	5
1.3. HIPÓTESIS DE TRABAJO.....	7
<b>CAPITULO II</b> .....	8
<b>REVISIÓN DE LA LITERATURA</b> .....	8
2.1. REVISIÓN DE LAS TEORÍAS DE DISCRIMINACIÓN LABORAL POR SECTOR ECONÓMICO (PÚBLICO Y PRIVADO) .....	8
2.2. REVISIÓN DE LA LITERATURA EMPÍRICA DE DISCRIMINACIÓN LABORAL POR SECTOR. ....	20
<b>CAPITULO III</b> .....	38
<b>DESCOMPOSICIÓN SALARIAL POR SECTOR</b> .....	38
3.1. DEFINICIÓN DEL MODELO DE ESCOLARIDAD – ECUACIONES MINCERIANAS. ....	38
3.1.1 MODELO DE ESCOLARIDAD Y TASAS DE RENDIMIENTO DE LA EDUCACIÓN .....	38
3.1.2 ECUACIONES MINCERIANAS .....	43
3.2. PRINCIPALES PROBLEMAS ECONOMETRÍCOS EN LA ESTIMACIÓN DE ECUACIONES MINCERIANAS.....	48
3.2.1 PROBLEMAS EMPÍRICOS EN LA ESTIMACIÓN DE ECUACIONES MINCERIANAS .....	48

3.2.2	CORRECCIÓN DE LOS PRINCIPALES PROBLEMAS ECONOMÉTRICOS EN LA ESTIMACIÓN DE ECUACIONES MINCERIANAS .....	56
3.3.	METODOLOGÍAS PARA LA DEFINICIÓN EMPÍRICA DE LOS FACTORES OBSERVABLES Y NO OBSERVABLES DE LOS DIFERENCIALES SALARIALES POR SECTOR .....	59
3.3.1	MODELOS ADITIVOS.....	60
3.3.2	MODELOS MULTIPLICATIVOS.....	62
3.3.3	DESCOMPOSICIÓN DE INGRESOS DE BLINDER-OAXACA.....	63
CAPITULO IV	.....	69
RESULTADOS	.....	69
4.1.	DESCRIPCIÓN DE LOS DATOS .....	69
4.2.	ANÁLISIS DESCRIPTIVO DE LOS DIFERENCIALES SALARIALES POR SECTOR LABORAL.....	77
4.3.	ESTIMACIONES: MODELOS ADITIVOS, MULTIPLICATIVOS Y DESCOMPOSICIÓN SALARIAL POR SECTOR.....	82
4.3.1	ANÁLISIS DE DISCRIMINACIÓN SALARIAL: MODELOS ADITIVOS Y MULTIPLICATIVOS.....	82
4.3.2	DESCOMPOSICIÓN SALARIAL POR SECTOR- “TWOFOOLD DECOMPOSITION BY SECTOR” .....	90
CAPITULO V	.....	95
CONCLUSIONES Y RECOMENDACIONES	.....	95
5.1.	CONCLUSIONES .....	95
5.2.	RECOMENDACIONES .....	98
BIBLIOGRAFÍA	.....	100

## LISTA DE GRÁFICOS

Gráfico- 1 Clasificación de la Población por Actividad.....	79
-------------------------------------------------------------	----



## LISTA DE TABLAS

Tabla- 1 Características de una Sub muestra de la ENEMDU por Sector Económico .....	81
Tabla- 2 Estimación de ecuaciones de ingresos poblacionales y por sector económico (Modelos aditivo, multiplicativo, público y privado).....	89
Tabla- 3 Descomposición de ingresos por sector “Twofold decomposition by sector”.....	94

## LISTA DE ANEXOS

ANEXO- A Validación del modelo multiplicativo.....	104
ANEXO- B Validación de los modelos por separado para los sectores público y privado.	122
Validación del modelo para el sector privado .....	122
Validación de modelo para el sector público .....	136

## RESUMEN

En el Ecuador se ha percibido desigualdad salarial entre sectores laborales; es por eso que el presente trabajo de investigación tiene como principal objetivo analizar empíricamente los diferenciales salariales entre el sector público y privado. Así, los resultados expuestos en este trabajo evidencian que efectivamente existe diferenciación salarial a favor del sector público (aproximadamente 0.50 puntos logarítmicos o 1.51 dólares por hora) y la mayor parte de esta diferencia se explica por las diferentes dotaciones entre grupos (52%), si se considera como estructura salarial no discriminatoria a la del sector público; pero si se considera como estructura salarial no discriminatoria a la del sector privado, la mayor parte de esta diferencia se debe a la discriminación salarial por sector laboral (57%). Estos datos estadísticos provienen de la estimación de ecuaciones mincerianas de ingresos por Mínimos Cuadrados Ordinarios (MCO) y por el método de Heckman (1979) que corrige el sesgo de selección muestral; además, se realiza la metodología de descomposición salarial propuesta por Blinder-Oaxaca (1973) que también es corregida por el método de Heckman(1979) y estimada por MCO.

**Palabras clave:** Sector público, diferenciación salarial, discriminación salarial, ecuación minceriana, descomposición de Oaxaca, Ecuador.

## ABSTRACT

In Ecuador has been perceived the wage inequality between employment sectors, that's why this research has as the principal aim, to analyze empirically the wage differentials between the public and private sectors. Thus, the results presented in this study show that there is indeed wage differentiation in favor of the public sector (about 0.50 log points or 1.51 U.S. dollars per hour) and most of this difference is explained by the different endowments between groups (52%), if one considers as non-discriminatory wage structure of the public sector, but if one considers as non-discriminatory wage structure of the private sector, most of this difference is due to the labor sector wage discrimination (57%). These statistical data come from the estimation of Mincer's earnings equations by Ordinary Least Squares (OLS) and by the method of Heckman (1979) to correct the sample selection bias; furthermore, is performed the wage decomposition methodology proposed by Blinder-Oaxaca (1973) which is also corrected by the method of Heckman (1979) and estimated by OLS.

**Keywords:** Public sector, wage differentiation, wage discrimination, Mincer's equation, Oaxaca decomposition, Ecuador.

# CAPITULO I

## INTRODUCCIÓN

En el Ecuador, aunque la incidencia de la pobreza ha presentado una tendencia decreciente (en el año 2006 el índice fue de 37.6% y para el 2011 disminuyó a 28.6%), y la línea de pobreza se ubicó en USD 2,40 diarios (2011), parecen existir todavía problemas de distribución del ingreso (en diciembre de 2011 el coeficiente de GINI fue de 0.473)<sup>1</sup> y de diferenciales salariales entre los diferentes grupos sociales y laborales.

En este sentido, para Carrillo (2004)<sup>2</sup>, en su estudio “Las diferencias salariales entre el sector público y privado en el Ecuador”, se demuestra la existencia de diferencias salariales entre estos dos sectores y se sugiere que en promedio los empleados públicos tienen salarios mayores en un 18% a sus equivalentes en el sector privado y las diferencias salariales varían fuertemente entre los asalariados que ganan menos, pero entre los asalariados que ganan más estas diferencias casi desaparecen; así las estadísticas muestran que dado un mismo nivel de educación, género, edad, estado civil, y origen étnico, el primer decil de la distribución de salarios de los trabajadores del sector público es 33% mayor al primer decil de la distribución de aquellos que trabajan en el sector privado; estas diferencias disminuyen conforme el salario se incrementa, pues en el noveno decil de la distribución, éstas ya no son estadísticamente significativas. En cuanto a las diferencias salariales entre sectores

---

<sup>1</sup>Los datos se obtuvieron de la página web del Instituto Nacional de estadísticas y Censos del Ecuador.

<sup>2</sup>Para conocer más de este estudio, véase Carrillo, Paul (2004), “Las diferencias salariales entre el sector público y privado en el Ecuador”, Cuestiones Económicas (Quito), Volumen 20, N. 2, pp. 165-173.

estratégicos tales como educación y salud, Carrillo (2004) muestra que el diferencial en el sector hospitalario es mayor al diferencial de la educación primaria; ya que los trabajadores hospitalarios (excluyendo médicos y odontólogos)<sup>3</sup> del sector público reciben remuneraciones con un 23% adicional en comparación a sus colegas del sector privado y los trabajadores de educación primaria pertenecientes al sector público reciben 10% más en sus salarios que los pertenecientes al sector privado; por lo tanto, según Carrillo (2004) estos resultados podrían evidenciar que las organizaciones de los empleados públicos han logrado incrementar los salarios por encima del nivel del mercado<sup>4</sup>. Cabe mencionar que en el Ecuador tanto el sector hospitalario como el de educación presionan frecuentemente por el incremento de sus salarios<sup>5</sup>. Por lo que, lo antes descrito puede evidenciar un problema, si se considera al diferencial salarial como un tipo de discriminación laboral que puede afectar a la pobreza e inestabilidad laboral; además, que los diferenciales salariales podrían contribuir a una distribución inequitativa de la riqueza en un país.

Por lo tanto, para poder resolver el problema identificado anteriormente el presente trabajo de investigación pretende analizar los diferenciales salariales entre el Sector Público y Privado para el período 2011 en el Ecuador, tomando en cuenta sus determinantes observables y no observables; los primeros, relacionados comúnmente desde la perspectiva empírica, con factores como la educación, experiencia, estado civil, localización geográfica, entre otros; y los segundos asociados con características no observables como la discriminación salarial o rendimientos.

---

<sup>3</sup>Se excluyen a médicos y odontólogos, tanto en el sector privado como el público, ya que además de las remuneraciones que perciben reciben ingresos adicionales, ya sea por concepto de honorarios por consultas privadas.

<sup>4</sup>El nivel de mercado para el trabajo de investigación de Carrillo (2004) es el ingreso mensual promedio de los asalariados de la submuestra de la EIGHU, que es de US\$ 370,63.

<sup>5</sup>Los resultados expuestos en el estudio de Carrillo (2004) sobre “Las diferencias salariales entre el sector público y privado en el Ecuador”, se basaron en la información de la Encuesta de Ingresos y Gastos de Hogares Urbanos del Instituto Nacional de estadísticas y Censos del Ecuador 2002-2003 (EIGHU).

Para cumplir con el objetivo antes mencionado se utilizará la metodología seguida por Torresano (2009)<sup>6</sup> en la que primeramente se estiman, ecuaciones mincerianas controladas por sector público y privado, en las cuales se analizan los signos y la significancia de sus coeficientes estimados, los que en caso de ser negativos, existiría la sospecha de discriminación salarial entre el sector público y privado. Sin embargo, este tipo de estimaciones son muy generales y restrictivas, ya que uno de los supuestos del modelo es que las estructuras de ingresos de todos los grupos son iguales<sup>7</sup>; por lo tanto, será necesario aplicar una descomposición de los diferenciales de ingresos por grupos aplicando la metodología de Blinder-Oaxaca (1973)<sup>8</sup>, quienes plantearon dicha descomposición de brechas salariales entre grupos en dos componentes: uno relacionado con las diferencias de las características observables de los grupos y otro atribuido a las diferencias en los rendimientos o características no observables relacionados con la discriminación salarial.

Es importante mencionar que el estudio que se va a realizar se fundamentará en las Bases de Datos del INEC de la Encuesta Nacional de Empleo, Desempleo y Subempleo (ENEMDU) del mes de diciembre de 2011; ya que según el INEC, dicha encuesta tiene como objetivo, “proporcionar información sobre el Mercado Laboral Ecuatoriano, a través de la recolección de datos, con periodicidad mensual en las principales ciudades y trimestral en el total nacional urbano, siguiendo las normas y recomendaciones internacionales de la OIT”. Entonces la ENEMDU facilita datos importantes para el estudio tales como salarios, sector público o privado al que pertenecen los trabajadores, actividad laboral, género, nivel de educación; en

---

<sup>6</sup>Torresano (2009) utiliza ecuaciones mincerianas controladas por género y etnia, además de modelos multiplicativos para interactuar las variables mencionadas con otras características observables como la región, estado civil, etc.; lo anterior con el objetivo de analizar los diferenciales salariales por género y etnia. Además utiliza la metodología propuesta por Oaxaca – Blinder (1973) para descomponer los diferenciales por género y etnia.

<sup>7</sup>El supuesto mencionado fue tomado de Torresano, Daniel (2009), “Análisis empírico de la discriminación salarial por género y etnia en el Ecuador en el año 2008”, Universidad Autónoma de Barcelona, Tesina, pp. 2.

<sup>8</sup>Blinder, Alan (1972), “Wage discrimination: reduced form and structural estimates”, *The Journal of Human resources*, Vol. 4, pp. 437-455. Y Oaxaca, Ronald (1973), “Male-Female wage differentials in urban labor markets”, *International Economic Review*, Vol. 14, No. 3, pp. 693-709.

conclusión, dicha encuesta proporciona características demográficas que pueden explicar las diferencias salariales entre los trabajadores.

A continuación se expone la estructura del presente trabajo de investigación. El primer capítulo hace referencia a la introducción que consta de los objetivos de la investigación, la justificación y la hipótesis del trabajo; en el segundo capítulo se resumen las principales teorías y desarrollos empíricos de discriminación laboral por sector; seguidamente, en el capítulo tres se explica la metodología a utilizarse para luego en base a los resultados de sus estimaciones efectuar el análisis empírico de discriminación salarial por sector laboral que se detalla en el capítulo cuatro, y finalmente, el capítulo cinco concluye y realiza las respectivas recomendaciones.

## **1.1. OBJETIVOS DE LA INVESTIGACIÓN**

### **1.1.1 OBJETIVO GENERAL**

Analizar empíricamente los diferenciales salariales entre el sector público y privado, mediante la descomposición de sus ingresos a través de la cuantificación de los determinantes observables y no observables para el Ecuador en el período 2011.

### **1.1.2 OBJETIVOS ESPECÍFICOS**

1. Identificar y cuantificar los determinantes observables y no observables de los diferenciales salariales entre el sector público y privado para el Ecuador.
2. Realizar la descomposición de ingresos entre el sector público y privado para el Ecuador en el período 2011.
3. Analizar empíricamente los diferenciales salariales entre el sector público y privado en el Ecuador, período 2011.



## 1.2. JUSTIFICACIÓN

### Justificación Teórica

El proyecto de tesis referente a los diferenciales salariales entre sector público y privado tiene gran importancia en el país, puesto que es un tema poco explorado, tanto desde la perspectiva teórica como empírica, ya que no se han encontrado muchos y significativos aportes a la literatura relacionada con el tema en análisis, lo que aparentemente no ha permitido el desarrollo científico o investigativo sustentado; por lo tanto, a través de esta investigación, se podría mejorar las condiciones para un análisis más consistente de las brechas salariales entre los grupos laborales en estudio. Además, desde la perspectiva empírica, se podrá cuantificar y conocer los posibles determinantes observables y no observables de los diferenciales salariales de los grupos, los cuales en posteriores análisis podrán servir de fundamento y referente en futuras investigaciones. Es importante mencionar que la aplicación de la metodología de Blinder-Oaxaca (1973) y sus refinaciones empíricas han sido muy poco aplicadas en el país por lo que el aporte científico en este tema ayudará al mejoramiento del fenómeno en estudio.

Adicionalmente, es un tema que con la información básica que se pueda obtener del Instituto Nacional de Estadísticas y Censos del Ecuador (INEC), y mediante la utilización de modelos econométricos como los que se mencionó en el marco introductorio se puede llegar a valiosas conclusiones, tales como: (i) la existencia o no de discriminación salarial entre el sector público y privado; y (ii) factores no observables y/o observables a los que se deben las diferencias salariales, entre otras. En el caso de que se encuentren diferencias salariales, no necesariamente esto significa que la eficiencia del mercado laboral esté distorsionada, ya que para Carrillo (2004) basándose en la teoría económica convencional manifiesta que si el sector privado ecuatoriano es competitivo, entonces el salario fijado por este, sería eficiente y equivalente a la productividad laboral de sus trabajadores, pero si el sector

público fija sueldos más altos, el equilibrio del mercado laboral se distorsionaría y como consecuencia el desempleo se incrementaría. En otro sentido, si el sector privado en el país tiene poder en el mercado, entonces los salarios fijados por este sector serían inferiores al de eficiencia y con salarios mejor remunerados por parte de la administración pública, existiría presión para eliminar las distorsiones, incrementando el bienestar de la sociedad. En conclusión, los resultados estadísticos podrán proporcionar información objetiva y técnica para la elaboración de políticas salariales tanto para el sector público como privado.

### **Justificación Práctica**

El desarrollo empírico descrito anteriormente, permitirá encontrar las causas de la discriminación salarial; las cuales suelen estar estrechamente relacionadas con la pobreza; por lo tanto, será de gran utilidad el estudio discriminatorio para encontrar alguna alternativa de solución tanto para la reducción de la pobreza como para las diferencias salariales; pues resolver estos problemas es de interés para organizaciones internacionales y el estado ecuatoriano.

En este sentido, existen varios métodos para medir la pobreza, siendo uno de ellos el de los ingresos. Este método es considerado en la presente investigación ya que la pobreza en un sentido amplio podría estar ligada con las diferencias salariales, considerando que la distribución de los individuos en cada nivel de pobreza podría ser alterado por la existencia de discriminación salarial; pues quienes con las mismas características de un individuo que percibe un mayor salario, al ser discriminado y por ende al obtener una menor remuneración, estará ubicado en un nivel de pobreza menor en comparación con ese individuo.

Otro aspecto que muestra la importancia del estudio, es que actualmente se ha podido percibir que el sector público está abarcando a una gran parte de la población económicamente activa; por ende la presente investigación podrá decir si esta afirmación es correcta, puesto que si el país cada vez está desembolsando más

dinero en remuneraciones para el sector público, con el transcurso del tiempo el presupuesto del estado se puede ver afectado e inclusive este puede caer en déficit.

### **1.3. HIPÓTESIS DE TRABAJO**

¿Entre el Sector Público y Privado en el Ecuador para el período 2011, existe discriminación salarial?

Dado que en el Ecuador en los últimos años ha existido transformaciones económicas, políticas y sociales con el fin de disminuir la pobreza, la verificación de la existencia o no de discriminación salarial entre los sectores público y privado puede ser de trascendental importancia para conseguir el objetivo antes mencionado; por lo tanto, la recolección de información, datos empíricos y la aplicación de diferentes metodologías econométricas y estadísticas son una de las mejores vías para la aceptación o rechazo de la hipótesis planteada. Por esta razón, para la prueba de la hipótesis mencionada en esta investigación, se utiliza la estimación de ecuaciones de ingreso y la denominada descomposición de los ingresos de los sectores público y privado del Ecuador para la cuantificación de los determinantes observables y no observables de los diferenciales salariales en el período 2011.

## **CAPITULO II**

### **REVISIÓN DE LA LITERATURA**

Previo a la revisión de la literatura tanto desde una perspectiva teórica como empírica de la discriminación salarial a continuación se considera una definición general de dicho concepto. Según Cain (1986), esta situación hace referencia al caso en el que el empleador considera no útil la contratación de un trabajador que pertenece a las minorías, únicamente por sus características demográficas y otras características observables, lo cual es irrelevante para la productividad física del trabajador, por lo tanto los posibles empleados son objeto de discriminación. En el desarrollo del presente capítulo se proporcionará una definición más formal de discriminación.

#### **2.1. REVISIÓN DE LAS TEORÍAS DE DISCRIMINACIÓN LABORAL POR SECTOR ECONÓMICO (PÚBLICO Y PRIVADO)**

Una de las primeras teorías de discriminación laboral es la Teoría Competitiva de las Diferencias de Grupos que para Altonji y Blank (1999) tiene tres elementos principales; el primero hace referencia a las Diferencias en Preferencias, que establece que las personas pueden seleccionar entre el mercado laboral (trabajo) o el mercado no laboral (ocio); así también pueden preferir entre tipos de trabajo particulares como la mano de obra versus el trabajo de oficina o trabajar en el sector público versus el trabajo en el sector privado. La distribución de las preferencias para

las características particulares del trabajo a través de grupos y el valor para los empleadores de ofrecer puestos de trabajo con características particulares determinarán la distribución de los salarios profesionales, así como la distribución ocupacional de grupos particulares.

El segundo elemento principal es las Diferencias en Ventajas Comparativas que manifiesta que en una economía competitiva estas diferencias influenciarán en la asignación del tiempo entre las ocupaciones y entre el mercado laboral y no laboral. Además, casi cualquier modelo de inversiones en capital humano muestra que las inversiones en habilidades valiosas será menor entre aquellos que esperan pasar menos tiempo en el mercado laboral. Un ejemplo de diferencias en ventajas comparativas es el de género; ya que en el pasado las mujeres empleaban sus habilidades en el cuidado y crianza de los hijos por lo que era menos probable que dichas mujeres se educaran para dedicarse a la medicina, ingeniería, abogacía, entre otras especialidades que requieren dedicar tiempo al mercado laboral, entonces la producción de sus habilidades era recompensado en el matrimonio, pues el hombre era quien se preparaba para el trabajo. Por lo tanto, si la familia juega un rol importante en la transmisión de preferencias para tipos de trabajo particulares y en la adquisición de una ocupación específica de capital humano, entonces históricamente determinadas diferencias en ventajas comparativas entre los grupos pueden persistir durante algún tiempo.

El último elemento de la teoría mencionada por Altonji y Blank (1999) es las Diferencias en Inversión en Capital Humano, donde aluden que el retorno de las habilidades generales adquiridas a través de la educación y la formación depende de la esperanza de participar en el mercado de trabajo, si estas habilidades incrementan la productividad del mercado laboral más que la productividad del mercado no laboral (ocio). Así también, el retorno para las firmas que invierten en capital humano y en investigaciones del mercado laboral es mayor si su personal trabaja a tiempo completo y no pretende dejar la empresa para participar en el mercado no laboral. Adicionalmente, hay una enorme literatura que documenta la importancia del historial

(background) familiar para los logros en educación y el éxito en el mercado laboral y principalmente la educación de los padres es la variable más importante en estos estudios.

Por otra parte, Altonji y Blank (1999) divide a los modelos económicos de discriminación en el mercado laboral en dos clases principales: los modelos competitivos en los cuales los agentes actúan individualmente y los modelos colectivos que enfatizan las consecuencias de la acción colectiva de un grupo contra otro, a menudo usando el sistema legal o la amenaza de la violencia como un mecanismo de aplicación. Estos dos modelos enfatizan dos tipos de discriminación; el primero es el prejuicio, el cual hace referencia al gusto de algunos miembros del grupo de la mayoría en contra de la interacción con los miembros del grupo minoritario. El segundo es una discriminación estadística por los empleadores en la presencia de información imperfecta acerca de las habilidades o comportamiento de los miembros del grupo de la minoría; en otras palabras para Altonji y Blank (1999), es una situación en la que los empleadores evalúan a los trabajadores usando limitada información proveniente de recomendaciones, entrevistas personales y hojas de vida; es decir, los empleadores a través de estos medios obtienen escasa información de los empleados. En tal sentido, la acumulación de experiencia laboral y la educación pueden ser los factores más importantes de la distribución de los ingresos entre trabajadores; también la antigüedad, la formación y la rotación del empleo son factores que causan las diferencias salariales entre grupos.

Para Cahuc y Zylberberg (2004), la discriminación estadística significa que individuos con idénticas habilidades, pero que pertenecen a grupos distintos no tienen trayectorias profesionales equivalentes a causa del promedio de cualidad, real o imaginaria del grupo del cual provienen. Además, añaden que la discriminación estadística es una fuente de la discriminación individual, pero no explica una discriminación entre grupos, ya que si los empleados son contratados en base a períodos de prueba o pruebas de contratación pueden no solo contratar a personas eficientes sino también a trabajadores ineficientes debido a que tienen la misma

probabilidad de pasar estas pruebas. Simultáneamente, Cahuc y Zylberberg (2004) mencionan que la discriminación estadística es una fuente de inequidad persistente entre grupos porque si se asume que la eficiencia de los trabajadores depende de la inversión en educación y en presencia de discriminación estadística, entonces el retorno de la educación es menor en la medida en que los empleadores creen que la proporción de trabajadores ineficientes en el grupo es considerable. Esta creencia puede actuar como un incentivo para los trabajadores para no educarse.

Como parte de la discriminación laboral existen diferentes grupos que tienden a trabajar en diferentes ocupaciones dando lugar a brechas salariales entre los mismos; a esto se lo denomina segregación laboral que puede surgir por muchas razones y para Altonji y Blank (1999) son tres: i) una posibilidad sería que los empleadores instauren discriminación en una ocupación particular, ii) otra alternativa es que los miembros de diferentes grupos seleccionan entre ocupaciones ya sea por las normas sociales, limitaciones legales o institucionales que restringen que los miembros de ciertos grupos puedan emplearse en algunas ocupaciones, además esta alternativa está de acuerdo que el fortalecimiento de los diferenciales salariales de los diferentes grupos puede ser causado por la acción colectiva y iii) la última causa es que las inversiones en capital humano antes de ingresar al mercado laboral o en actividades fuera del mercado laboral pueden diferir entre grupos y conducir a diferencias en las ventajas comparativas entre ocupaciones.

Por otra parte, para Altonji y Blank (1999) una definición teórica más formal de discriminación establece que ésta es una situación en la cual personas quienes proveen de servicios en el mercado laboral, que son igualmente productivas en un sentido físico o material son tratadas en forma desigual dada una característica observable como la raza, etnia, género, o sector económico. En particular la desigualdad observada desde la perspectiva económica es la de los salarios.

Para Becker (1973) existen tres tipos de discriminación que son: discriminación del empleador, discriminación del empleado y discriminación de los consumidores. Detallando a la discriminación del empleador como una situación en la cual los

empleadores tienen perjuicios en contra de los miembros del grupo de la minoría. Por lo demás, los empleadores maximizan su utilidad que es la suma de los beneficios más el valor monetario de la utilidad de emplear a los miembros de un grupo en particular. Así también, describe a la discriminación del empleado como el escenario donde algunos miembros del grupo de la mayoría tienen perjuicios contra los miembros del grupo de la minoría y no les gusta trabajar con este último grupo. Supone además que existen dos tipos de trabajadores que son los trabajadores calificados y los no calificados, de igual forma existen dos tipos de trabajo, el calificado y el no calificado y en las ocupaciones no calificadas todos los empleados tienen el mismo nivel de productividad. Finalmente, define a la discriminación de los consumidores al caso en que los consumidores del grupo de la mayoría obtienen una utilidad menor (considerándose perjudicados) si ellos compran a los miembros del grupo de la minoría; pues obtendrían una utilidad mayor si su elección sería comprar a miembros del grupo mayoritario. Consecuentemente, el grupo de la mayoría comprará al de la minoría solo si el precio de venta es menor, reduciendo la rentabilidad del mercado de trabajo para el grupo de la minoría que trabajan en ocupaciones en contacto con clientes.

Cahuc y Zylberberg (2004) ofrecen una definición de discriminación como una situación en la que individuos idénticos en lo que respecta su habilidad productiva son tratados de forma diferente debido a algunas de sus características no productivas. Además mencionan que la discriminación de empleados y empleadores no puede ocurrir en un mercado perfectamente competitivo ya que por definición todos los trabajadores son pagados de acuerdo a su productividad marginal; en consecuencia, la discriminación está enlazada a la competencia imperfecta como por ejemplo el poder de los monopsonios puede implicar discriminación entre grupos con habilidades productivas que son previamente idénticas.

Adicionalmente, Becker (1973) alude que los científicos sociales definen a la discriminación del mercado de acuerdo a sus disciplinas; así señala que los Sociólogos indican que los diferentes niveles de discriminación en contra de un grupo



en particular están asociados con diferentes niveles sociales, distancias físicas de ese grupo o con diferentes niveles de estatus socioeconómico. Además, Becker (1973) da una ilustración de discriminación de mercado contra un grupo minoritario, la cual se da cuando la discriminación ha reducido la tasa media salarial neta del grupo minoritario por un mayor porcentaje que el grupo mayoritario. Del mismo modo, Becker (1973) señala que la discriminación está comúnmente asociada con la des-utilidad causada por contacto con algunos individuos y además dicha discriminación incorpora tanto el prejuicio y la ignorancia de la verdadera eficiencia económica de un grupo de personas; sin embargo, la ignorancia puede ser rápidamente eliminada por la difusión del conocimiento; mientras que un prejuicio (o preferencia) es relativamente independiente del conocimiento. Igualmente, la discriminación se relaciona con determinantes geográficos y la localización cronológica, porque la discriminación puede variar de país a país, de región a región dentro de un país, de áreas rurales a áreas urbanas dentro de una región y de un período de tiempo a otro. Finalmente, el gusto por discriminación puede diferir simplemente debido a las diferencias en la personalidad de los miembros de los grupos (discriminado y discriminador).

Posteriormente, Becker (1973) menciona que el dinero sirve como una medida de discriminación e indica que si un individuo tiene algún gusto por la discriminación, él preferiría relacionarse con algunas personas en lugar de otras y en consecuencia estaría dispuesto a realizar un pago directo o un pago en forma de ingreso reducido para complacer su gusto, del mismo modo si un individuo tiene gusto por la discriminación, él debe actuar como si estuviera dispuesto a perder ingresos con el fin de evitar ciertas transacciones.

Analíticamente, Becker (1973) se basa en los coeficientes de discriminación para expresar la definición de gusto por la discriminación. Por lo tanto, bajo el supuesto de que existen dos grupos en el mercado laboral designados por  $W$  y  $N$  y sean los miembros de  $W$  perfectamente sustitutos en la producción para los miembros de  $N$ , el coeficiente de discriminación del mercado (CDM) es:

$$CDM = \frac{\pi_w - \pi_n}{\pi_n} \quad (1)$$

Donde  $\pi_w$  y  $\pi_n$  son las tasas salariales de equilibrio para  $W$  y  $N$  respectivamente, cabe recalcar que en ausencia de discriminación y favoritismo y si el mercado laboral es perfectamente competitivo, el salario de equilibrio de  $W$  sería igual al de  $N$ ; pero en presencia de discriminación, existiría diferencia en los salarios. Por lo tanto el coeficiente de discriminación del mercado es determinado como la diferencia proporcional entre estas tasas salariales ya sea en el caso en que  $W$  y  $N$  sean sustitutos imperfectos<sup>9</sup>, entonces el CDM es igual a la diferencia entre la relación de las tasas salariales de  $W$  y  $N$  con o sin discriminación, siendo representado de la siguiente manera:

$$CDM = \frac{\pi_w}{\pi_n} - \frac{\pi_w^0}{\pi_n^0} \quad (2)$$

Sean  $\pi_w^0$  y  $\pi_n^0$  son las tasas salariales de equilibrio para  $W$  y  $N$  respectivamente sin discriminación y  $\pi_w$  y  $\pi_n$  son las tasas salariales de equilibrio para  $W$  y  $N$  en presencia de discriminación.

Becker (1973) determina coeficientes de discriminación individuales (CD) tanto para los empleadores, empleados y consumidores. Los CD representan para el empleador un costo o retorno no monetario de producción, para el empleado un costo o retorno no monetario de empleo y para el consumidor es un costo o retorno no monetario de consumo; además los CD constituyen un puente entre el dinero y los costos netos. Así, suponiendo que un empleador se enfrenta con una tasa de salario  $\pi$  de un factor particular; él está actuando como si la tasa de salario neta fuera  $\pi (1+d_i)$ , con  $d_i$  como el CD en contra de este factor. En el caso de los empleados, ellos ofrecen una tasa salarial  $\pi$  para trabajar con ese factor y actúan como si la tasa salarial neta fuera  $\pi (1-d_j)$ , con  $d_j$  como el CD en contra de este factor. En cuanto a los consumidores, ellos se enfrentan con un precio unitario monetario ( $p$ ) para los bienes producidos por

---

<sup>9</sup>Según Becker (1973), si los grupos  $W$  y  $N$  son sustitutos imperfectos pueden recibir tasas salariales diferentes incluso en ausencia de discriminación.

este factor y actúan como si el precio neto fuera  $p(1+d_k)$ , con  $d_k$  como el CD en contra de este factor. Las cantidades  $\pi d_i$ ,  $\pi d_j$  y  $\pi d_k$  son los equivalentes monetarios exactos de los costos (o retornos) no monetarios para determinadas tasas salariales y precios. En estos casos el CD es la proporción que permite el cambio de los costos monetarios (o los retornos de dinero) expresados en dinero a magnitudes netas; además permite estimar su costo neto salarial; por lo tanto el coeficiente de discriminación es un elemento no pecuniario en ciertos tipos de transacciones y es positivo o negativo dependiendo si este elemento es considerado bueno o malo. Aditivamente, el coeficiente de discriminación de un empleador frente a un empleado mide el valor del costo no pecuniario de emplearlo, ya que representa la diferencia porcentual entre el dinero y tasa salarial neta pagada a los empleados. El coeficiente de discriminación de otros empleados contra estos empleados mide el valor del costo no pecuniario de trabajar con ellos y representa la diferencia porcentual entre el dinero y tasa salarial neta recibida por trabajar con esos empleados y el coeficiente de discriminación de un consumidor mide el valor del costo no pecuniario de comprar un producto producido por el grupo minoritario y representa la diferencia porcentual entre el dinero y el precio neto pagado por ese producto.

Paralelamente, Becker (1973) muestra en forma analítica que la discriminación de los empleadores convierte la tasa salarial monetaria del grupo N ( $\pi_n$ ) en una tasa salarial neta  $\pi_n(1+d_a)$ , donde  $d_a$  representa el CD de los empleadores. La discriminación directa del Gobierno contra los miembros del grupo N incrementan los costos no salariales de emplear a N y aumenta la tasa salarial neta a  $\pi_n(1+d_a+d_b)$ , donde  $d_b$  mide el costo incurrido. Ahora si los miembros del grupo W, se asocian y excluyen a los miembros del grupo N a causa de la discriminación contra ellos, y si los sindicatos puede imponer los costos no salariales en las empresas que emplean a los miembros del grupo N, la tasa salarial neta aumenta a  $\pi_n(1+d_a+d_b+d_c)$ , donde  $d_c$  mide el costo impuesto por el sindicato o la alianza mencionada. En consecuencia la tasa salarial neta total es equivalente a  $\pi_n(1+d_1)$ , siendo  $d_1 = d_a + d_b + d_c$  que representa la combinación de los efectos de la discriminación por parte de los empleadores, gobernantes y sindicatos.

Otro de los conceptos relacionados a la discriminación es el de Discriminación Económica que según Cain (1986) es definida en términos de diferencias de ingresos entre familias y diferencias salariales entre trabajadores. En cuanto a la discriminación en el mercado laboral, este es medido por las disparidades de los salarios y hace referencia al caso en el que el empleador considera no útil la contratación de un trabajador que pertenece a las minorías, únicamente por sus características demográficas y otras características observables, lo cual es irrelevante para la productividad física del trabajador, por lo tanto los posibles empleados son objeto de discriminación.

Cain (1986), señala que para el análisis de la discriminación económica es posible basarse en dos unidades de observación como son: i) la familia, la cual es la unidad apropiada para examinar las diferencias de bienestar económico entre los grupos mayoritario y minoritario, usualmente medido por el ingreso y ii) el trabajador que es la unidad apropiada para examinar las desigualdades en las tasas salariales o de ganancias. Adicionalmente, los ingresos monetarios anuales son asumidos como el indicador de bienestar económico y la diferencia será expresada como una relación del ingreso del grupo de la minoría con el del grupo mayoritario.

Cada una de las dos unidades de observación (familia y trabajador) pueden ser analizadas mediante un modelo lineal aditivo, el cual es recomendado para aplicaciones en el corto plazo y analíticamente es representado por Cain (1986) de la siguiente manera:

$$Y_i = BX_i + AZ_i + e_i \quad (3)$$

Donde:

$Y_i$ = resultado del proceso como el ingreso, salario o ganancias del individuo  $i$ ,

$X_i$ = variable exógena que no depende de  $Y$  y es un vector de características productivas del individuo  $i$ , tales como: experiencia potencial<sup>10</sup>, años de educación, etc.,

$Z_i$ = variable dicotómica discreta igual a 1 si el individuo pertenece al grupo de la minoría y cero si pertenece al grupo de la mayoría, ambos grupos proporcionan servicios laborales iguales, condicionados a  $X$ ,

$A$  y  $B$ = vector de coeficientes<sup>11</sup>, y

$e_i$ = término de error.

Si en una regresión  $A > 0$  podría evidenciarse discriminación y también  $A$  puede ser considerada como la discriminación del mercado ( $D$ )<sup>12</sup>.

Un segundo modelo que Cain (1986) plantea, representa al ingreso así:

$$Y = CZ + \mu \quad (4)$$

Donde  $\mu$  es el término de error y si  $C > 0$  hay evidencia de discriminación. El presente modelo puede ser aplicado para el largo plazo y asume que todos los trabajadores son igualmente productivos; no obstante la desventaja del mismo está en que ignora la distinción común entre la ocurrencia de discriminación dentro y la ocurrencia de discriminación previo o fuera del mercado laboral.

Según Cain (1986), la diferencia entre el primer y segundo modelo planteado anteriormente radica en que el problema práctico de disparidades del bienestar económico, usualmente definido en términos de diferencias de ingreso familiar, es

---

<sup>10</sup>Según Altonji y Blank (1999), la experiencia potencial para cada individuo puede ser calculada como el resultado de la diferencia de la edad - los años de educación - 5 y bajo el supuesto de que las personas están trabajando en su edad adulta cuando no están en la escuela.

<sup>11</sup> $A$  y  $B$  son coeficientes que representan los efectos de las variables  $Z$  y  $X$  respectivamente, sobre la variable dependiente  $Y$ .

<sup>12</sup>El coeficiente  $A$  es considerado como la discriminación del mercado ( $D$ ), ya que sea  $\hat{y}$  el valor condicional predicho de  $Y$  sobre  $X$ , la discriminación del mercado está dada por la siguiente ecuación:

$$D = (\hat{Y}|X, Z = 1) - (\hat{Y}|X, Z = 0)$$

dirigido por el segundo modelo y el práctico y teórico problema de diferencias en los salarios en trabajadores con igual productividad es examinado por el primer modelo, aunque el primer modelo puede ser especificado lo más cerca posible al segundo modelo como se desee mediante la restricción del conjunto de variables características admisibles.

Torresano (2009) manifiesta que Cain (1986) reconoce que los modelos mencionados tienen ciertas limitaciones relacionadas con su definición como por ejemplo la medida y significado del ingreso como un índice de bienestar económico, ya que si se toma en cuenta a la medición del salario como un precio, es necesario diferenciar entre retornos corrientes y permanentes o retornos pecuniarios y no pecuniarios, también en el tiempo, con la medida y evaluación del ocio y la recompensa del trabajo del hogar. Además, Torresano (2009) también menciona que para Altonji y Blank (1999) existe otra complicación en la definición de la ecuación (3)<sup>13</sup> que está relacionada con la distinción entre discriminación en el mercado laboral corriente dado un conjunto predeterminado de características observadas de los trabajadores y los efectos de la discriminación previa sobre estas características. Este último caso se puede entender si se considera por ejemplo, la existencia de discriminación en el interior del hogar o en el acceso a la educación en una generación inicial, lo que podría generar diferentes efectos como la reducción del nivel corriente de educación en un grupo minoritario, pero que en general pueden ser denominados como efectos de pre discriminación en el mercado laboral. A pesar de las limitaciones mencionadas, la ecuación (3) presentada por Cain (1986) es una herramienta altamente utilizada para la determinación de discriminación salarial entre los diferentes grupos de trabajadores.

Por otra parte, para Cain (1986) las diferencias salariales pueden ser una medida fundamental de la discriminación del mercado laboral. Una primera justificación a lo mencionado es que cuando la discriminación toma la forma de negación

---

<sup>13</sup>La dificultad de la definición de la ecuación (3) para Altonji y Blank (1999) no solo está relacionada con la distinción entre discriminación en el mercado laboral corriente y los efectos de la discriminación previa, sino que también está vinculada con la definición de igualdad de productividades individuales.

generalizada a contratar o promover a los trabajadores que pertenecen al grupo de la minoría, esto debería hacer que los salarios relativos bajen y una segunda justificación es que la discriminación salarial puede existir independientemente del grado de integración o segregación en el mercado.

Para Cahuc y Zylberberg (2004), las diferencias salariales están enlazadas con la competencia individual, explicadas por la teoría del capital humano que se basa en la idea que la educación conduce a los individuos a ser competentes de tal forma que tengan valor en el mercado laboral. También las diferencias salariales se relacionan con la teoría hedónica de los salarios que a continuación se presenta.

La predicción de la teoría hedónica de los salarios ilustra características individuales observables y no observables, al igual que preferencias heterogéneas de los individuos acerca de las características atractivas y no atractivas de algún trabajo. El modelo de predicción de la teoría hedónica de los salarios según Cahuc y Zylberberg (2004) es de la forma siguiente:

$$\ln(w) = X\beta + \alpha e + \varepsilon \quad (5)$$

Donde:

$w$ = salario recibido por un individuo,

$X$ = vector de características personales que generalmente incluye la edad, sexo, nivel de educación, experiencia, antigüedad en el trabajo, origen étnico, lugar de residencia, estatus familiar y afiliación sindical,

$e$ = vector de características no salariales del trabajo, incorpora variables como la duración y flexibilidad de horas de trabajo, el aspecto repetitivo de tareas, el riesgo de lesiones, el nivel de ruido ambiental, la fuerza física requerida para el trabajo, el riesgo de perder el trabajo, el costo incurrido en salud, el costo de ahorro para la jubilación, etc.,

$\alpha$  y  $\beta$ = vectores de parámetros estimados, y

$\varepsilon$ = término de perturbación con cero significa que sigue una distribución normal.

En resumen, el presente trabajo de investigación pretende entender la discriminación sectorial mediante un análisis empírico de los diferenciales salariales entre el sector público y privado; por lo tanto para lograr dicho objetivo se utiliza el modelo aditivo (3) planteado por Cain (1986); ya que la ecuación explica los salarios en función de variables productivas de los trabajadores (experiencia potencial, nivel de educación, etc.) y distingue mediante una variable binaria a los grupos minoritarios y mayoritarios que ofrecen servicios laborales idénticos pero que se encuentran condicionados a las características productivas. A pesar de la utilidad del modelo para entender la discriminación sectorial, este presenta ciertas limitaciones descritas en párrafos anteriores que serán tomadas en cuenta y que se intentará corregirlas desde una perspectiva empírica particularmente a través de metodologías estadísticas y econométricas que serán detallados en capítulos posteriores.

## **2.2. REVISIÓN DE LA LITERATURA EMPÍRICA DE DISCRIMINACIÓN LABORAL POR SECTOR.**

El propósito de este estudio es cuantificar las diferencias salariales entre la administración pública y el sector privado; por consiguiente es necesario utilizar métodos estadísticos convencionales que se presentan en la literatura de la economía laboral. Además, alrededor del mundo, de Latinoamérica y en el Ecuador existen análisis empíricos de discriminación laboral por sector que generalmente utilizan una metodología común, la cual se mencionara a continuación.

Al observar la literatura empírica aplicada al fenómeno de estudio, se puede apreciar primeramente que la gran mayoría de estudios realizan estimaciones de ecuaciones mincerianas controladas por sector, además introducen variables de productividad individual (como experiencia, nivel de educación, etc.). Por otra parte, para el análisis



de dichas ecuaciones, se verifica la significancia y los signos de sus coeficientes, ya que si son negativos podría evidenciarse discriminación salarial (o premio sectorial). Posterior a las estimaciones de las ecuaciones mincerianas en toda la población (tanto para individuos trabajadores del sector público y privado) y con el objetivo de corregir ciertos problemas de sesgo, se efectúa la descomposición de Oaxaca de brechas salariales por grupos. Dicho procedimiento inicialmente realizado por Blinder-Oaxaca (1973) propone una desagregación de dichas brechas en dos componentes: **un observable** atribuido a las diferencias en las dotaciones de los grupos y un componente **no observable** atribuido a las diferencias en las características no observables relacionado con la discriminación salarial. Por último, algunos estudios aplican un análisis de sensibilidad para corregir problemas de sesgo de selección muestral, a través de la aplicación del método propuesto por Heckman (1979). Es importante mencionar que todas las ecuaciones planteadas son estimadas inicialmente por el método de Mínimos Cuadrados Ordinarios (MCO) para posteriormente realizar las correcciones antes mencionadas.

Después de conocer el procedimiento econométrico que siguen ciertos trabajos de discriminación salarial, se describe brevemente algunos estudios referentes al tema.

En relación a la literatura empírica internacional, Heitmueller (2006) realiza un importante aporte empírico que tiene como objetivo determinar si existe una prima de ingresos para un escocés que pertenece al sector público y cómo los resultados difieren en comparación con otros estudios en el Reino Unido.

Heitmueller (2006) utilizó la base de datos del Estudio de Panel de Hogares Británico (BHPS) para analizar las diferencias salariales entre el sector público y privado en Escocia para el año 2000. Además, la muestra<sup>14</sup> de empleados que utiliza contiene a los individuos que son a la fecha de la entrevista empleados a tiempo completo y su edad es de 16 a 64 años (de 16 a 59 años para las mujeres), no trabajan por cuenta

---

<sup>14</sup>La sección transversal de la muestra se compone de 1.054 hombres y 1.230 mujeres de los cuales 61% de los varones y el 42% de las mujeres están participando en el mercado de trabajo. De éstos, aproximadamente el 22% de los hombres y el 41% de las mujeres están empleadas en el sector público.

propia y no trabajan ya sea en la agricultura, las organizaciones no lucrativas o para las fuerzas armadas.

Por otra parte, la metodología que utiliza Heitmueller (2006) es una estimación por Mínimos Cuadrados Ordinarios (MCO) de ecuaciones de ingresos por separado tanto para el sector público como para el sector privado; sin embargo Heitmueller (2006) reconoce que pueden existir estimaciones inconsistentes, pues las estimaciones de MCO son propensas a tener problemas de sesgo de selección de la muestra debido a la exclusión de los no participantes en la fuerza laboral; asimismo es posible que la asignación al empleo público o privado no sea aleatoria. En consecuencia, se realiza la corrección de sesgo de selección utilizando el método econométrico sugerido por Heckman (1979) y finalmente se aplica la descomposición de Neuman y Oaxaca (2002).

Las variables utilizadas en el trabajo de Heitmueller (2006) son: el ingreso por hora de los trabajadores del sector público y privado (variable dependiente), variables explicativas como características de los individuos (edad, estado civil, educación, puesto de trabajo, pertenencia a sindicatos) y variables latentes (dependientes), una para la elección de participar o no en el mercado laboral y otra para la elección de pertenecer al sector público o privado.

Los resultados de las estimaciones econométricas realizadas por Heitmueller (2006) muestran que el logaritmo de los salarios medios de los hombres en el sector público y privado son 2,2 y 2,0 respectivamente, por consiguiente, la diferencia en las medias es de 0.2132 puntos logarítmicos con el 5% de nivel de significancia, lo que revela que las remuneraciones del sector público son significativamente mayores que las remuneraciones del sector privado. De la misma manera, para el caso de las mujeres, los salarios en el sector público son mayores que los del sector privado, pues el logaritmo del salario medio en el sector público es de 2.15 y en el sector privado es de 1.74; (la diferencia de los salarios medios es de 0.4140 puntos logarítmicos con el 5% de nivel de significancia); por lo tanto existe una brecha

salarial entre los sectores público y privado tanto para los hombres como para las mujeres.

Adicionalmente a los resultados anteriores, Heitmueller (2006) identifica que el porcentaje de la brecha salarial está asociada a las características observables y a las no observables para el caso de los hombres y mujeres por separado. En relación al género masculino y tomando en cuenta al sector privado como estructura salarial no discriminatoria, el 106 % de las diferencias salariales se deben a las características observables (tales como estado civil, edad, nivel de educación, pertenencia a sindicatos, puesto de trabajo)<sup>15</sup> y parte de la brecha de ingreso atribuida a estas características es de 0.23 puntos logarítmicos (significativa al 5%); pero si la estructura salarial no discriminatoria es la del sector público dicho componente representa el 104% de los diferenciales salariales por sector y la parte de la brecha de ingreso es de 0.22 puntos logarítmicos (significativa al 5%). Con respecto a la contribución de las características no observables (como la discriminación salarial o rendimientos) en los diferenciales salariales tanto para el sector público y privado como estructuras salariales no discriminatorias es negativa y no significativa.

En cuanto al género femenino, si la estructura salarial no discriminatoria es la del sector privado, el 46% de la diferencia salarial se explica por las diferencias en las características observables ya mencionadas en el párrafo anterior, además la parte de la brecha de ingreso atribuida a estas características es de 0.19 puntos logarítmicos (significativa al 5%); pero, si la estructura salarial no discriminatoria es la del sector público, la diferencia salarial se explica por las diferencias en las características observables en un 92%; adicionalmente, la parte de la brecha de ingreso atribuida a estas características es de 0.38 puntos logarítmicos (significativa al 5%). Haciendo referencia al componente atribuido a la discriminación

---

<sup>15</sup> Para observar resultados relacionados con las diferencias de los ingresos respecto del estado civil, edad, o demás características observables; también para conocer resultados relacionados con las variables que determinan la elección de pertenecer al mercado laboral o al sector público y privado, véase Heitmueller, Axel (2006), "Public-Private Sector Pay Differentials In A Devolved Scotland", *Journal of Applied Economics*, Vol IX, No. 2, pp. 295-323.

(características no observables), este aporta con el 8% en las diferencias de ingresos tomando en cuenta como grupo minoritario al sector público. De modo similar para el sector privado como grupo minoritario, el componente atribuido a la discriminación contribuye el 54% en el diferencial salarial; además la parte de la brecha de ingreso es de 0.22 puntos logarítmicos (significativa al 5%).

Otro trabajo que hace referencia a la literatura empírica internacional de discriminación laboral por sector es el de Albert y Moreno (1998) el cual tiene como principal objetivo analizar las diferencias salariales entre el sector público y privado en España. En este trabajo se investiga si las diferencias son consecuencia de la distinta composición de empleo entre los sectores; también se examina si la elección de trabajar en el sector público o privado está determinada significativamente por el diferencial salarial.

Para la aplicación de los modelos econométricos que se describen más adelante, Albert et al. (1998) utilizan la base de datos de la Encuesta de Conciencia y Biografía de Clases (ECBC) que proporciona las características personales y laborales de los entrevistados; además esta base está compuesta por asalariados y no asalariados<sup>16</sup>.

El procedimiento metodológico aplicado en Albert et al. (1998) implicó la estimación la ecuación de salarios de Mincer (1974) para estudiar las diferencias salariales, tomando en cuenta al logaritmo neperiano del salario neto mensual como variable dependiente y como variables independientes el logaritmo neperiano del número de horas trabajadas, años de escolarización, edad; además se incluye una variable dicotómica que identifica si el trabajador pertenece al sector público o privado. Seguidamente, se estiman dos ecuaciones de salario distintas, una para cada sector; sin embargo, se aclara que las estimaciones pueden tener problemas de autoselección (que consiste en que la decisión de trabajar en uno u otro sector, no es

---

<sup>16</sup> Supone 6632 casos; además, Albert et al. (1998) indican que el diseño de la muestra se realizó a partir del Censo Electoral para el total del estado pero sobre representando a la Comunidad de Madrid; por consiguiente, las estimaciones se efectúan sobre los datos ponderados con el objetivo de corregir la sobre representación de la Comunidad Autónoma de Madrid y del colectivo de los individuos con estudios superiores.

una variable exógena sino una variable endógena); en consecuencia, para evitar que las estimaciones sean sesgadas los autores del trabajo aplican el método en etapas de Heckman que no sólo tiene en cuenta el problema de endogeneidad de trabajar en el sector público o privado, sino que también intenta corregir el problema de selección previa de ser asalariado o no. Posteriormente, Albert et al. (1998) realizan una descomposición de las diferencias salariales en tres factores que son: a) las diferencias originadas por la composición del empleo en los dos sectores, es decir, por las diferencias en las características de los trabajadores de cada sector, b) las diferencias originadas por los distintos coeficientes que reflejan las estructuras de pago de cada sector, es decir, por la diferente remuneración que tienen las características en cada sector y c) las diferencias salariales debidas a los sesgos de selección. Además, es importante mencionar que Albert et al. (1998) aplican una ecuación general de descomposición salarial, la cual es la media entre dos ecuaciones de descomposición; la primera ecuación toma como estructura salarial no discriminatoria la del sector privado y la segunda ecuación toma como estructura salarial no discriminatoria la del sector público; pues la estructura salarial no debe de elegirse arbitrariamente y debe fundamentarse en la teoría.

Luego de la aplicación de los modelos econométricos corregidos por sesgo de selección, Albert et al. (1998) en general concluyen que los salarios en términos medios son mayores en el sector público en comparación con las remuneraciones del sector privado; es decir los resultados muestran que cuando se define como sector público a la Administración y Empresa Pública la diferencia media entre los salarios de los sectores antes mencionados es de 16% en contra del sector privado, porcentaje menor a cuando se toma la definición de sector público exclusivamente a la Administración Pública ya que su diferencial salarial medio es del 23% a favor del sector público. Adicionalmente, la diferencia salarial entre ambos sectores ocasionada por las diferencias entre las medias de las características de los mismos es del 28% en contra del sector privado, si se considera como parte del sector público a la Administración y Empresa Pública, pero si se define únicamente a la Administración Pública como sector público, la diferencia salarial entre sectores

originada por la composición del empleo es del 27% en contra del sector privado. Así también, las estimaciones indican que una misma característica es mejor remunerada en el sector privado que en el público, ya que para el caso de la Administración y Empresa Pública como parte del sector público, la diferencia salarial media entre el sector público y privado debido al componente no observable (rendimientos) es del 14% a favor del sector privado (existe un penalidad para el sector público). Igualmente, para el caso en que solo la Administración Pública es considerada como Sector Público, la diferencia salarial media entre el sector público y privado debido a los rendimientos es del 17% a favor del sector privado (al igual que para la definición anterior existe un premio para el sector privado). Finalmente, los valores de las diferencias salariales a favor del sector público debidas a los sesgos de selección son 12% y 1.6% a favor del sector público para el caso de que la Administración Pública sea considerada sector público y para el caso de que la Administración y Empresa Pública sean parte del sector público respectivamente.

También, en base a las estimaciones de Albert et al. (1998) se puede evidenciar la contribución de las características observables y no observables al diferencial salarial entre sectores. Por lo tanto, si se considera a la Administración y Empresa Pública como parte del sector público las características observables representan el 175% de las diferencias salariales totales y los sesgos de selección explican el 10% del diferencial salarial, pero si solo la Administración Pública es considerada como sector público las características observables explican el 117% de las diferencias salariales totales y los sesgos de selección contribuyen el 52% al diferencial salarial. Por último, la representación de las características no observables en el diferencial salarial es negativa (para las dos definiciones de Sector Público).<sup>17</sup>

---

<sup>17</sup>Adicionalmente a estos resultados, Albert et al. (1998) estudian las diferencias entre los sectores privado y público en base a las características del empleo (horas de trabajo, antigüedad, nivel de educación, entre otros) y descubren que las horas medias trabajadas son ligeramente inferiores en el sector público; conjuntamente revelan que un incremento en las horas trabajadas tiene mayor repercusión sobre los salarios en el sector privado que en el público. Además, para conocer cómo las variables (edad, estado civil, genero, etc.) y las diferencias salariales influyen en la decisión de pertenecer o no al mercado laboral y en la elección de pertenecer al sector público o privado, el lector puede dirigirse a Albert, Cecilia y Moreno, Gloria (1998), "Diferencias salariales entre el sector

En relación a la literatura empírica latinoamericana, Holanda et al. (2012) investigan las diferencias salariales entre los trabajadores de los sectores público y privado en Brasil. Por otra parte, las ecuaciones econométricas estimadas se basan en los micro datos de la Encuesta Nacional de Hogares (PNAD) de 2009, realizada por el Instituto Brasileño de Geografía y Estadística (IBGE). Además, el universo de análisis abarca todas las personas de 18 a 65 años de edad, los residentes de las zonas urbanas que reportaron una semana de trabajo de entre 20 y 72 horas y un retorno positivo en su ocupación principal durante la semana de referencia<sup>18</sup>.

Las variables que Holanda et al. (2012) utilizan en sus modelos econométricos comprenden al logaritmo neperiano del salario como variable endógena y entre las variables exógenas se encuentran variables demográficas tales como raza, género, variables productivas como nivel de educación (representado por las variables primaria, media y superior, correspondientes a los años de escolaridad completados en cada nivel de educación) y experiencia<sup>19</sup> (definida como la edad del trabajador menos los años de escolaridad y menos seis) y una variable que identifica la ubicación geográfica del trabajador.

Específicamente, el modelo de identificación de la posibilidad de trabajar en el sector público o privado contiene variables relacionadas con las características demográficas y productivas ya mencionadas, adicionalmente se incluyen variables de las restricciones de exclusión que pueden ayudar en la identificación del problema de la elección de sector, tales como el número de hijos del trabajador de entre 0 y 15 años, el número de personas mayores en la familia, la condición de ser cabeza de familia y las rentas no provenientes del trabajo.

---

público y privado español: un modelo de switching”, Estadística Española, Vol. 40, Núm. 143, 1998, pp. 167 a 193.

<sup>18</sup> Holanda et al. (2012) indican que la muestra utilizada se restringió a 85.526 personas (alrededor del 66% de la población original reportado por PNAD).

<sup>19</sup> Esta medida incorpora el supuesto de que todos los trabajadores comienzan su vida escolar a los seis años de edad y están trabajando en su edad adulta cuando no están en la escuela. También, la experiencia se incluye en su forma cuadrática ( $EXPERIENCIA^2$ ) con el fin de tener en cuenta la posible depreciación de capital humano.

Con respecto a la metodología econométrica empleada, Holanda et al. (2012) definen un modelo aditivo de salarios tanto para el sector público como para el sector privado bajo el supuesto de que los trabajadores están distribuidos al azar entre ambos sectores. Además, el análisis de estas dos ecuaciones compara los salarios observados, de esta manera, los sectores público y privado sólo cubren a los trabajadores que ya están empleados en uno de ellos, y no tienen en cuenta el hecho de que los trabajadores eligieron (o fueron elegidos) para trabajar en cada sector. Por tanto, existe un sesgo de autoselección ya que los trabajadores pueden emplearse en cualquier sector de forma no aleatoria; en consecuencia, los modelos generan estimaciones sesgadas e inconsistentes.

Para la corrección del problema de auto selección, Holanda et al. (2012) aplican un modelo de regresión econométrica con cambio endógeno que se realiza en dos etapas. En la primera etapa, se estima un modelo probit con el fin de determinar aquellas variables que tienen efecto sobre la probabilidad de entrar en el sector público o privado. En la segunda etapa, la Inversa del Ratio Mills para cada sector se añade en las dos ecuaciones de salarios para la corrección del sesgo de selección<sup>20</sup>. En conclusión, este modelo intenta controlar cualquier ventaja o desventaja comparativa que los dos grupos tienen en sus respectivos sectores.

Es importante destacar que las estimaciones de las ecuaciones de salarios se realizan a través del método de mínimos cuadrados ordinarios (MCO) y que el diferencial de los sectores público y privado se calcula según el género, en la medida en que la participación en el mercado de trabajo es muy diferente entre hombres y mujeres. Además, Holanda et al. (2012) no aplican la descomposición de Oaxaca.

Los resultados en el estudio de Holanda et al. (2012) muestran que en Brasil los salarios del sector público son más altos que los del sector privado<sup>21</sup>. El cálculo de la brecha salarial se realiza entre las medias de los ingresos nominales de los

---

<sup>20</sup> La inclusión de la inversa del ratio de Mills en las ecuaciones salariales toma en cuenta la decisión del trabajador de elegir entre el sector público y el sector privado.

<sup>21</sup> Adicionalmente a estos resultados Holanda et al. (2012) indican los factores que influyen en la probabilidad de que un individuo que trabaja en el sector público.



dos sectores, por tanto, el salario promedio del sector público es superior en un 84% con respecto al salario promedio del sector privado.<sup>22</sup> También, Holanda et al. (2012) muestran la diferencia logarítmica de las medias de los salarios estandarizados en base al modelo corregido por sesgo de selección; por lo tanto la diferencia es de 0.1763 puntos logarítmicos a favor del sector público y significativa al nivel del 5%.

Adicionalmente, Holanda et al. (2012) indican que la conclusión de que los salarios del sector público son mayores a los salarios del sector privado también se fundamenta en que los gastos de personal del gobierno consolidado representó casi el 13% del Producto Interno Bruto (PIB) en 2007.

En cuanto al análisis específico por género, se observa que, el diferencial salarial estandarizado para el género femenino entre los sectores público-privado alcanza el 108%, mientras que para los hombres es de 127%, ambos diferenciales a favor del sector público. Así también, los resultados del modelo corregido indican que el diferencial logarítmico del promedio de los salarios estandarizados es de 0.0172 en contra del sector privado para el caso del género femenino y para el género masculino este diferencial es de 0.1138 puntos logarítmicos, igualmente, en contra del sector privado (ambos diferenciales significativos al 5%).

Asimismo, entre los trabajos que hacen referencia a la literatura empírica Latinoamérica, **Panizza (2000)**, analiza las diferencias salariales entre los sectores público-privado y a su vez analiza las brechas de género para América Latina; por consiguiente utiliza los datos de las encuestas de hogares que abarca 17 países de América Latina durante el período 1981-1998. Además, a fin de mitigar el sesgo de selección debido a la asistencia a la escuela, este trabajo toma en cuenta a las personas de entre 20 y 65 años de edad.

---

<sup>22</sup> Parte del diferencial del rendimiento estandarizado se explica por la diferencia de horas de trabajo entre el sector público y el sector privado. De hecho, es importante señalar que el promedio de las horas de trabajado semanal en el sector público es de 37,81 que es casi 13% menos que el promedio de las horas semanales trabajadas en el sector privado, pues su media es de 43,49 horas.

En dicho estudio, la variable dependiente utilizada fue el logaritmo natural del ingreso por hora y las variables de control relacionadas con las características productivas de los individuos (como experiencia, experiencia al cuadrado, 5 variables dicotómicas de educación, 8 variables binarias del sector de ocupación y una variable dicotómica para el área de residencia); también se incluye una variable binaria que toma el valor de uno para las mujeres trabajadoras y una variable dicotómica que toma el valor de uno si el trabajador es empleado en el sector público.

La metodología que Panizza (2000) utiliza consiste en estimar por Mínimos Cuadrados Ordinarios (MCO) varias ecuaciones. Primeramente plantea dos ecuaciones de ingresos, una para estimar el premio del sector público y otra para estimar la brecha de género; después estima un modelo multiplicativo donde existe una interacción entre variables binarias de género y sector público. Seguidamente, Panizza (2000) intenta resolver problemas de estimaciones sesgadas; por consiguiente, para solucionar el problema del supuesto de que los hombres y las mujeres empleados en los sectores público y privado tienen la misma estructura de rendimientos, realiza la estimación de ecuaciones mincerianas de ingresos para cada uno de los grupos, por lo tanto, estima ecuaciones por separado para hombres y mujeres y ecuaciones por separado para los sectores público y privado; también estima ecuaciones separadas para los trabajadores con niveles altos y bajos de educación. Posteriormente, el supuesto de las dos primeras ecuaciones de que la decisión de participar en el mercado de trabajo y la clasificación de los trabajadores entre los sectores público y privado es aleatoria, no necesariamente puede cumplirse; en consecuencia, Panizza (2000) sugiere un análisis de sensibilidad. A pesar de estas restricciones, Panizza (2000) asegura que las estimaciones de MCO de las dos primeras ecuaciones y del modelo multiplicativo darán lugar a una medida bastante exacta del diferencial salarial del sector público-privado y la brecha de género en América Latina. Además, es importante mencionar que Panizza (2000) no realiza la descomposición de Oaxaca.

En base a las estimaciones que Panizza (2000) obtuvo en la mayoría de los países de América Latina, no son grandes las diferencias entre los salarios de los trabajadores varones que pertenecen a los sectores público y privado. En general, en promedio, los datos indican una pequeña prima del sector público para los trabajadores varones, 2% para toda la muestra y 4% cuando se considera a los trabajadores poco cualificados. También, si se considera las diferencias salariales entre los sectores público y privado para el caso femenino, el sector público paga una prima significativa a las mujeres que oscila entre el 9% y 32% (especialmente a aquellas con pocas habilidades). Particularmente, tanto para el género masculino y femenino, las mayores penalidades en el sector público corresponden a Bolivia y República Dominicana (significativas al 1% y la mayor penalidad es 40%) y por el contrario Costa Rica, Colombia, Ecuador y El Salvador reciben las mayores primas por parte del sector público.

En cuanto a la brecha de género, en el sector público es de aproximadamente 10%, 20 puntos porcentuales por debajo de la brecha de género del sector privado; adicionalmente, la brecha antes mencionada es bastante homogénea en todos los niveles de educación<sup>23</sup>, ya que en el sector privado esta es significativamente mayor (aproximadamente 8 puntos porcentuales) cuando sólo se consideran los trabajadores poco calificados. Específicamente, este documento indica que las mujeres ocupadas en el sector privado tienen una brecha de género de 17 puntos porcentuales (en promedio) más que la brecha de género del sector público; además, las mujeres latinoamericanas que trabajan en el sector privado ganan un 30% menos que los hombres con habilidades similares, y aproximadamente una tercera parte de esta brecha de género se debe a la falta de acceso al sector formal. Con respecto a los países destacados en el sector privado, la brecha de género es mayor en Honduras con el 43% y menor en Colombia con el 20%, ambas significativas al 1%.

---

<sup>23</sup> Los resultados muestran que en América Latina en promedio un 16% de las personas empleadas trabajan para el sector público y más del 30% de los trabajadores con educación superior son empleados en este sector. Además, las estadísticas indican que el 40% de los hombres y el 60% de las mujeres que trabajan en el sector formal son empleados por el sector público. El sector público absorbe cerca del 60% de los trabajadores que poseen títulos universitarios o han aprobado la escuela secundaria.

La situación es diferente cuando nos fijamos en el sector público, ya que las brechas de género son elevadas (hasta un 36 por ciento) y significativas para Bolivia, Brasil, Costa Rica, Chile, Colombia, Panamá, Uruguay y Venezuela.

Por otra parte, el diferencial de discriminación<sup>24</sup> se encuentra alrededor del 50 por ciento para Guatemala, Honduras y República Dominicana y en menos del 8 por ciento para Brasil. Además, el diferencial es positivo en 14 de los 17 países de América Latina (significativo en 12 de estos países).

Otro estudio de América Latina, lo efectúan posteriormente **Panizza y Zhen-Wei Qiang (2005)** y su fin es investigar las diferencias salariales entre los sectores público y privado en 13 países de América Latina; también analizan cómo el sector del empleo afecta a la brecha salarial de género.

Este documento utiliza una colección de las encuestas de hogares de 13 países de América Latina<sup>25</sup>, que cubren más del 80% de la población total de la región y permite identificar si un trabajador es empleado en el sector público o en el sector privado. Además, al igual que el estudio anterior para mitigar el sesgo de selección debido a la asistencia a la escuela, la investigación sólo incluye en el análisis a las personas de 20-65 años edad y como punto de referencia para los empleados del sector público, toma en cuenta a los trabajadores empleados en el sector formal (definido como las empresas del sector público más las empresas del sector privado con más de 6 empleados) y excluye a los trabajadores autónomos y a los militares.

Las variables que Panizza et al. (2005) toman en cuenta en su estudio son: una variable dependiente que es el logaritmo natural del salario por hora en el trabajo

---

<sup>24</sup> Diferencial entre la brecha de género de los sectores público y privado.

<sup>25</sup> Los países de América Latina de los cuales se analizan las diferencias salariales entre los sectores público y privado son: Bolivia (año de encuesta 1997), Brasil (año de encuesta 1995), Colombia (año de encuesta 1997), Costa Rica (año de encuesta 1995), El Salvador (año de encuesta 1995), Ecuador (año de encuesta 1995), Honduras (año de encuesta 1996), México (año de encuesta 1994), Nicaragua (año de encuesta 1993), Panamá (año de encuesta 1995), Paraguay (año de encuesta 1995), Perú (año de encuesta 1997) y Venezuela (año de encuesta 1995). Además todas las encuestas son a nivel nacional a excepción de Colombia, la cual provee de una encuesta urbana.

principal de cada individuo, variables de control que representan características individuales (como experiencia, experiencia al cuadrado, 5 variables dicotómicas de educación), una variable binaria que toma el valor de uno para las mujeres trabajadoras, otra variable binaria que toma el valor de uno si el trabajador es empleado en el sector público y una interacción entre las variables dicotómicas de género y sector público. También están incluidas variables dicotómicas del sector de ocupación (minería, manufactura, construcción, agua y electricidad, comercio minorista, restaurantes y hoteles, transporte y telecomunicaciones, servicios financieros, y otros servicios) y área de residencia (diferenciando áreas rurales y urbanas).

Panizza et al. (2005) realizan estimaciones utilizando el método de Mínimos Cuadrados Ordinarios (MCO) y el procedimiento metodológico que se describe a continuación.

En un primer momento se realiza una estimación de un modelo multiplicativo de ingresos, controlado por sector público, sector privado y género, además este modelo involucra la interacción entre sector público y género; después se estima la ecuación de salarios controlada por sector; dado que este modelo presenta ciertas complicaciones econométricas como el supuesto de que los hombres y mujeres tienen el mismo conjunto de beneficios a sus características individuales, Panizza et al. (2005) reconocen que se debería realizar una segunda estimación por MCO que es la de dos etapas desarrollado por Heckman (1979), sin embargo, no la realizan ya que no encontraron una variable que sea correlacionada con el sector de empleo y que no afecte a los ingresos del trabajador; en consecuencia, mediante la realización de simulaciones asumen diferentes niveles de correlación entre el salario y las regresiones de selección; no obstante, aunque ya no se impone en este modelo el mismo conjunto de beneficios para hombres y mujeres, existe el supuesto de que el vector de coeficientes de las variables explicativas es el mismo para los trabajadores empleados en los sectores público y privado; por lo tanto, se aplica el modelo propuesto por Blinder y Oaxaca (1973) para descomponer las diferencias salariales

mediante la estimación de ecuaciones diferentes para los grupos que deben ser comparados. Finalmente, Panizza et al. (2005) aplican un análisis de sensibilidad, pues el supuesto en el que se basa la descomposición de Oaxaca de que los trabajadores están distribuidos al azar entre los sectores público y privado puede no cumplirse.

En resumen, este trabajo realiza una aportación metodológica que consiste en probar la robustez de los resultados de la estimación de MCO de las diferencias salariales de los sectores público y privado, bajo distintos supuestos para la correlación entre la ecuación que determina el sector del empleo (público vs privado) y la ecuación de salarios.

Por otra parte, los resultados del documento señalan que en la mayoría de los países latinoamericanos, los empleados del sector público disfrutan de una prima salarial sobre los empleados del sector privado. También encuentra que la prima del sector público tiende a ser mayor para las mujeres que para los hombres.

El trabajo además descubre considerables diferencias de género en los sectores público y privado. Con respecto a la brecha de género, a las mujeres se les paga menos que a los hombres en ambos sectores, sin embargo, los datos sugieren que en la mayoría de los países la brecha de género es mayor en el sector privado. Cuantitativamente, las estimaciones por MCO indican una brecha de género promedio de 10% para el sector público y 16% para el sector privado<sup>26</sup>.

También, Panizza et al. (2005) para el caso del género masculino explica en base a sus resultados que los salarios promedios de los trabajadores son siempre más altos en el sector público, el promedio no ponderado es del 57%. Además, los resultados de la descomposición de Oaxaca<sup>27</sup> bajo el supuesto de que la correlación entre la ecuación de salarios y las regresiones de selección es cero y utilizando la

---

<sup>26</sup> Para observar detalladamente resultados referentes a la brecha de género véase Panizza, Ugo y Zhen-Wei Qiang, Christine (2005), "Public-private wage differential and gender gap in Latin America: spoiled bureaucrats and exploited women?", *Journal of Socio-Economics*, Vol. 34, pp. 810-833.

<sup>27</sup> Panizza et al. (2005) también muestra los resultados de las estimaciones del modelo Cotton-Neumark.

rentabilidad del sector privado como punto de referencia, sugieren que entre los países de América Latina existen varias diferencias de los salarios medios, así por ejemplo los trabajadores de Bolivia del sector público reciben en promedio 6% más en sus salarios en relación a las remuneraciones del sector privado, los trabajadores públicos de Venezuela reciben un salario mayor en 10% con respecto a los trabajadores privados, en Ecuador los trabajadores varones del sector público reciben un salario adicional del 43% en comparación con los trabajadores varones del sector privado, adicionalmente, existen dos casos relevantes en las diferencias de salarios entre sectores ya que los trabajadores varones que pertenecen al sector público de Brasil y Colombia reciben más del 100% en sus remuneraciones con respecto al sector privado. La mayor parte de esta diferencia se explica por el hecho de que, en promedio, los trabajadores del sector público tienden a una dotación superior ya que las diferencias medias intersectoriales de las características observables de los trabajadores es de 29%, 13% y 37% a favor del sector público para Bolivia, Venezuela y Ecuador respectivamente; para Brasil es de 176% y para Colombia 78% en ambos casos a favor del género masculino perteneciente al sector público. Adicionalmente, en ocho de los 13 países, parte de las diferencias salariales se debe a una prima del sector público que van desde el 4% (Brasil) a 23% (Colombia). En cinco países, se encontró con una penalidad del sector público, que van desde el 3% (Venezuela y Panamá) al 18% (Bolivia).

Al igual que en el caso del género masculino, Panizza et al. (2005) encuentran grandes diferencias salariales entre las mujeres empleadas en el sector público y las mujeres empleadas en el sector privado. La diferencia salarial media es más alta que la de la muestra de trabajadores varones ya que para las mujeres esta es del 64% frente a la de los varones que es del 57%. Igualmente, para la interpretación de las estimaciones se recurre a la descomposición de Oaxaca asumiendo que la correlación entre el salario y las regresiones de selección es cero y utilizando la rentabilidad del sector privado como punto de referencia, por lo tanto, los resultados indican que la diferencia salarial entre los sectores público y privado para el género femenino oscilan entre el 6% (Venezuela) y más del 100% (191% para Brasil, 117%

para El Salvador y 115% para Colombia); en el caso ecuatoriano, la diferencia salarial entre los sectores público y privado es de 30% a favor del sector público. De forma similar al caso de los hombres, el diferencial se debe principalmente a las diferencias en las características observables, cuyos valores varían desde un mínimo que es 4% para Venezuela hasta un valor máximo que corresponde a Brasil con el 185% y específicamente para el Ecuador es del 28%. En relación al premio o penalidad del sector público para las mujeres trabajadoras, los valores varían entre -22% (Bolivia) y 37% (El Salvador) y para Ecuador es del 2%.

Para el caso ecuatoriano aparentemente solo **Carrillo (2004)** ha estudiado el fenómeno de discriminación salarial; dicha investigación cuantificó las diferencias salariales entre los sectores público y sector privado.

Carrillo (2004) en su estudio ha utilizado la Encuesta de Ingresos y Gastos de Hogares Urbanos 2002-2003 (EIGHU) del Instituto Ecuatoriano de Estadísticas y Censos (INEC); pero para las estimaciones eligió una submuestra de individuos de la EIGHU que consta de 2812 asalariados con las siguientes características: a) ser asalariado del sector formal, b) trabajar por lo menos 30 horas a la semana, y c) no pertenecer a una organización cuya actividad esté vinculada con la Agricultura, Extracción de Petróleo, Manufactura, y Generación Eléctrica.

El estudio mencionado define una ecuación semi-logarítmica para explicar los determinantes del salario y su estimación lo hace por medio de Mínimos Cuadrados Ordinarios (MCO); posteriormente utiliza el Método de Regresión de Cuantiles para identificar las diferencias salariales en cada cuantil de la distribución salarial.

Entre las variables que utiliza Carrillo (2004) están como variable dependiente el salario por hora y como variables explicativas se tomó en cuenta al género, edad, años de educación, nivel de educación, etnia, estado civil y una variable dummy que toma el valor de uno si el individuo trabaja en el sector público.

Los resultados que Carrillo (2004) obtuvo sugieren que en promedio los empleados públicos tienen salarios mayores en un 18% a sus colegas en el sector privado.



Además, las estadísticas muestran que dado un mismo nivel de educación, género, edad, estado civil, y origen étnico, el primer decil de la distribución de salarios de los trabajadores del sector público es 33% mayor al primer decil de la distribución de aquellos que trabajan en el sector privado; estas diferencias disminuyen conforme el salario se incrementa. En cuanto a las diferencias salariales entre sectores estratégicos tales como educación y salud, Carrillo (2004), muestra que los trabajadores hospitalarios del sector público<sup>28</sup> reciben remuneraciones con un 23% adicional en comparación a sus equivalentes del sector privado y los trabajadores de educación primaria pertenecientes al sector público reciben 10% más en sus salarios que los pertenecientes al sector privado.

En cuanto a las diferencias salariales entre hombres y mujeres, las estimaciones sugieren que en promedio las mujeres ganan un salario menor en 8.2% a relación de los hombres, evidenciando los problemas de discriminación de género en el mercado laboral ecuatoriano.

Por otra parte, para Carrillo (2004) las estimaciones de las diferencias de salario debido al nivel de escolaridad evidencian que por un año adicional de educación el salario se incrementa en aproximadamente 7%, mientras que aquellos que han realizado estudios de postgrado, en promedio, perciben un salario mayor en 29% más a relación de aquellos que no lo han hecho. Los resultados también sugieren que el salario se incrementa con la edad pero con una tasa decreciente.

En conclusión, los salarios del sector público son en promedio 17.5% más altos que las remuneraciones del sector privado.

---

<sup>28</sup> Excluyendo médicos y odontólogos porque además de las remuneraciones que perciben, reciben ingresos adicionales, ya sea por concepto de honorarios por consultas privadas.

## **CAPITULO III**

### **DESCOMPOSICIÓN SALARIAL POR SECTOR**

#### **3.1. DEFINICIÓN DEL MODELO DE ESCOLARIDAD – ECUACIONES MINCERIANAS**

Previo a la definición de las ecuaciones mincerianas, es necesario detallar el modelo de escolaridad y las tasas de rendimiento de la educación ya que son la base para la derivación de las ecuaciones mencionadas; por consiguiente, a continuación se expone los fundamentos requeridos.

##### **3.1.1 MODELO DE ESCOLARIDAD Y TASAS DE RENDIMIENTO DE LA EDUCACIÓN**

Para la descomposición de los diferenciales salariales por sector público y privado, se requiere la estimación de ecuaciones mincerianas. Este procedimiento facilita la determinación de los rendimientos medios de la educación. En este sentido, a continuación se detalla en forma sucinta la teoría que se encuentra detrás de la estimación de las ecuaciones mincerianas y su relación con la economía de la educación.

El modelo de escolaridad propuesto por Mincer (1974) calcula los efectos de la escolarización sobre las ganancias. Willis (1986) explica que el modelo mencionado considera una economía hipotética compuesta por trabajadores que se diferencian por tres elementos que son:

- i) Años de escolaridad,  $s$  (que se supone que empiezan a la edad de 6 años),
- ii) Edad,  $t$ ,  $y$
- iii) Duración de la experiencia<sup>29</sup> de la fuerza laboral,  $x = t - s - 6$ .

Por lo tanto, en esta economía, para Willis (1986), los datos sobre los ingresos anuales ( $y$ ), los años de escolaridad ( $s$ ) y la experiencia ( $x$ ) pueden ser descritos por una función<sup>30</sup> estadística de ingresos que se muestra a continuación:

$$y = \varphi(s, x) + u, \quad (6)$$

Donde  $\varphi(s, x)$  es la función que mejor se ajusta a los datos y  $u$  es un residuo con media cero que capta el efecto de las variables no observadas tal como la capacidad que influyen en la productividad individual. Además, la ecuación (6) se presenta bajo el supuesto de que  $u$  es estadísticamente independiente de  $s$  y  $x$ .

Por otra parte, Willis (1986) menciona que otro supuesto del modelo es que el cambio de la productividad agregada es cero y la tasa de crecimiento de la población es constante con una distribución estable de edades asociado.

También, para poder aplicar las ecuaciones mincerianas es necesario entender su descomposición para lo cual se requiere conocer la definición de la tasa marginal de rendimiento interno que según Willis (1986) es la tasa de descuento,  $p(s_1, s_2)$ , tal que el valor presente de los flujos de los ingresos netos de los costos directos de la educación que están asociados a dos diferentes niveles de escolaridad,  $s_1$  y  $s_2$ , se equiparan. Para el cálculo del valor presente existen supuestos tales como: a) el único costo de la escolaridad son los ingresos no percibidos, b) los individuos ingresan a la fuerza laboral inmediatamente después de la finalización de la escuela

<sup>29</sup> Para Willis (1986), el perfil de los ingresos con respecto a la experiencia es de pendiente positiva en la mayoría o toda la vida laboral y cóncava con la tasa de crecimiento de las ganancias siendo mayor en edades tempranas y disminuyendo o incluso tornándose negativa en etapas posteriores de la vida.

<sup>30</sup> Según Willis (1986) la ecuación (1) es interpretada como una función hedónica de precios en el sentido de Rosen (1974), que refleja el equilibrio de la oferta y la demanda de trabajadores en cada nivel de la educación y la experiencia.

a la edad  $t = 6 + s$ , y c) la vida laboral de cada individuo de  $n$  años es independiente de sus años de educación.

Además, Willis (1986) indica que teniendo en cuenta un supuesto adicional de un estado de equilibrio sin crecimiento de la productividad, el valor presente de los ingresos durante la vida de un individuo "representante" con  $s$  años de educación, evaluada a la edad de ingreso a la escuela, es:

$$V(s, r) = \int_0^n \varphi(s, x) e^{-r(s+x)} dx, \quad (7)$$

Siendo  $r$  es la tasa de descuento y  $\varphi(s, x)$  se basa en la función estadística estimada ganancias (6) y  $n$  los años de vida del individuo.

Ahora, Willis (1986) plantea dos niveles de escolaridad distintos; por lo tanto, sea  $d > 0$  y  $\hat{p}(s, s+d)$  una aproximación de la tasa interna marginal de retorno de un individuo con  $s$  años de escolarización que invierte en  $d$  años adicionales de instrucción. Por definición,  $\hat{p}(s, s+d)$  es la tasa de descuento que soluciona  $V(s, r) = V(s+d, r)$ . Utilizando (7), se puede demostrar que esta definición implica que:

$$\hat{p}(s, s+d) = \frac{1}{d} \{ \ln(\int_0^n \varphi(s+d, x) e^{-r((s+d)+x)} dx) - \ln(\int_0^n \varphi(s, x) e^{-r(s+x)} dx) \} \quad (8)$$

Por otra parte, si se considera el supuesto de que la tasa de crecimiento de las ganancias en cualquier nivel de experiencia dado es independiente del mismo; es decir, por ejemplo un incremento dado de la educación tiene el mismo efecto proporcional en los ingresos de todos los niveles de experiencia. En este caso, Willis (1986) rescribe la función de ganancias (6) de la siguiente manera:

$$y = f(s)g(x) + u \quad (9)$$

Lo que implica que el rendimiento estimado a un aumento pequeño en la escolarización por encima de un nivel dado de  $s$  es:

$$\ln(y) = \ln(f(s)) + \ln(g(x)) + \ln(u)$$

$$\frac{d \ln y}{ds} = \frac{f'(s)}{f(s)} = \hat{p}(s) \quad (10)$$

Donde  $\hat{p}(s)$  es una aproximación de la tasa marginal de rendimiento interno de la educación. A continuación se muestra su justificación:

El valor presente de los ingresos de toda la vida cambia así:

$$V(s, r) = f(s)e^{-rs} \int_0^n g(x)e^{-rx} dx \quad (11)$$

Además, en este caso, la tasa marginal interna de retorno estimada a la educación está dada por la derivada del logaritmo de la función de rentabilidad estadística (9) con respecto a  $s$ . Por lo tanto, usando (8) - (10), se sigue que:

$$\hat{p}(s, s + d) = [\ln(f(s + d)) - \ln(f(s))]/d \quad (12)$$

Si el supuesto mencionado en (9) no se cumple, Mincer (1974) ha sugerido un "atajo", un método aproximado para calcular la tasa de retorno a la educación<sup>31</sup>.

El método de atajo de Mincer implica el uso de un concepto de adelantamiento y supone que los ingresos medios evolucionan según la función de ingresos  $y = \varphi(s, x)$  y  $V_s$  es el valor presente del perfil de ganancias. Además, si se considera a  $\bar{y}(s)$  un nivel constante de ganancias, las cuales tienen el mismo valor presente y  $x^*(s)$  el nivel de experiencia de adelantamiento, tal que,  $\bar{y}(s) = \varphi(s, x^*(s))$  implique  $p(s, s + d) = [\ln(\varphi(s + d, x^*(s + d))) - \ln(\varphi(s, x^*(s)))]/d$ . En este sentido, se tiene que el valor presente del perfil de ganancias se expresa mediante la siguiente ecuación:

$$V_s = \int_0^n \bar{y}(s)e^{-rx} dx$$

$$V_s = \alpha \bar{y}(s)/r$$

$$V_s = \alpha \varphi(s, x^*(s))/r \quad (13)$$

---

<sup>31</sup> El método de atajo sugerido por Mincer (1974) evita la necesidad de utilizar métodos numéricos y tiene la ventaja de que la tasa de rendimiento puede estimarse a partir de los datos de los primeros diez años de historia de una cohorte de ganancias.

Donde  $\alpha=(1-e^{-rn})$  y  $x^*(s+d)$  es el nivel de experiencia de adelantamiento para el perfil de ganancias  $\varphi(s+d,x)$  que está asociado con un mayor nivel de estudios,  $s+d$ . Sustituyendo (13) en (8), la tasa interna de retorno marginal es:

$$p(s, s + d) = \frac{1}{d} \{ \ln[(1 - e^{-rn})\varphi(s + d, x^*(s + d))] * (1/r) ] - \ln[(1 - e^{-rn})\varphi(s, x^*(s))](1/r) \}$$

$$p(s, s + d) = \frac{1}{d} \{ \ln \left[ \frac{(1 - e^{-rn})}{r} \right] + \ln[\varphi(s + d, x^*(s + d))] - \ln \left[ \frac{(1 - e^{-rn})}{r} \right] - \ln[\varphi(s, x^*(s))] \}$$

$$p(s, s + d) = [\ln(\varphi(s + d, x^*(s + d))) - \ln(\varphi(s, x^*(s)))]/d \quad (14)$$

Además, Willis (1986) manifiesta que para Mincer (1974) el nivel de experiencia de adelantamiento será menor que o igual a la inversa de la tasa interna de retorno. Análogamente, Willis (1986) indica que Mincer (1974) también desarrolla el argumento de adelantamiento para un caso especial en el que el nivel de experiencia a adelantar es exactamente  $1/p$ , independientemente del nivel de escolaridad. Así también, supone que los individuos ingresan a la fuerza laboral con una capacidad de ingresos de  $\bar{y}(s)$  dólares e invierten  $C$  dólares en el puesto de trabajo en cada año después de salir de la escuela por el que pagan  $C$  dólares de los ingresos no percibidos durante el período de inversión, también, las inversiones tienen una tasa de retorno constante  $p\%$  a perpetuidad. Dados estos supuestos, los ingresos de un individuo con  $s$  años de años de escolaridad y  $x$  años de experiencia es:

$$y(s, x) = \bar{y}(s) + p \int_0^x C dt - C = \bar{y}(s) + pCx - C = \bar{y}(s) + C(px - 1) \quad (15)$$

Bajo el supuesto de que el trabajador tiene una vida infinitamente larga de trabajo y la tasa de descuento es  $p$ , entonces el valor actual del flujo de ganancias en la ecuación (15) es  $\bar{y}(s)/p$  que también corresponde al valor presente de un flujo de ingresos constante de  $\bar{y}(s)$ . Simultáneamente es posible verificar que a partir de la ecuación (15)  $x^*(s) = 1/p$ , ya que es el valor de  $x$  que resuelve  $y(s, x) = \bar{y}(s)$ .

Empíricamente, Willis (1986) manifiesta que en la ecuación (15) el crecimiento en dólares de los ingresos tiende a disminuir a medida que aumenta la experiencia. En consecuencia, el nivel constante de ganancias con el mismo valor presente como  $y(s, x)$  tendería a ser menor que  $\bar{y}(s)$  y el punto de adelantamiento tendería a ocurrir antes. Por lo tanto, para Mincer el nivel de experiencia de adelantamiento tenderá a ser algo menor que  $1/\rho$ .

En particular, Willis (1986) da cuenta que las personas con altas tasas de inversión en formación en el puesto de trabajo tienden a tener menores ganancias iniciales y mayor crecimiento de las ganancias en comparación con los individuos que invierten a una tasa menor. Así, la varianza de las ganancias entre los individuos tienden a tener forma de U con el mínimo en  $x=1/\rho$ , suponiendo que la capacidad de ingresos inicial y la tasa de inversión después de la escuela no están correlacionadas.

### 3.1.2 ECUACIONES MINCERIANAS

En el presente apartado se realiza un breve resumen de la teoría económica que sustenta la estimación de las denominadas ecuaciones mincerianas. Por tanto, para Willis (1986), la ecuación minceriana básica es:

$$\ln(y) = \beta_0 + \beta_1 s + \beta_2 x + \beta_3 x^2 + u \quad (16)$$

Donde,  $y$  son los ingresos anuales,  $s$  los años de escolaridad y  $x$  los años de experiencia.

De la interpretación de esta ecuación se desprende que los incrementos porcentuales de los ingresos son estrictamente proporcionales a las diferencias absolutas del tiempo invertido en educación, con la tasa de retorno como el coeficiente de proporcionalidad. Más precisamente, la ecuación minceriana muestra el logaritmo de las ganancias como una función lineal del tiempo empleado en educación.

Por otra parte, Willis (1986) describe la función de ganancias del capital humano desarrollado por Mincer (1974) bajo el supuesto de que los trabajadores continúan invirtiendo en capacitación, pero a una tasa decreciente, después de su período de escolaridad. Adicionalmente, si se considera la ecuación (12), se puede mostrar que el coeficiente de escolarización  $\beta_1$  de (16) proporciona una estimación de la tasa de retorno a la educación<sup>32</sup>. Además, la ecuación (16) tiene que ser cóncava<sup>33</sup>, en consecuencia  $\beta_3 < 0$ .

Willis (1986) también manifiesta que la justificación de Mincer para la función de ganancias del capital humano consiste en una mezcla de la teoría y el pragmatismo. En cuanto a la **teoría**, se supone que las habilidades (o competencias) adquiridas por los trabajadores a través de la educación y en el puesto de trabajo pueden ser consideradas como un stock de capital humano homogéneo que influye en la productividad del trabajador en la misma cantidad en todas las líneas de trabajo para todos los empleadores<sup>34</sup>.

En relación a la inversión en la educación Willis (1986) explica que si la empresa ofrece capacitación, implícitamente, se cargará al trabajador mediante la reducción de sus salarios por debajo de su producto marginal por el coste de la formación y los trabajadores están dispuestos a pagar este precio implícito debido al aumento de sus ingresos futuros derivados de su mayor productividad. Por lo tanto, los ingresos observados de un trabajador a un determinado nivel de experiencia pueden ser considerados como iguales a la tasa de alquiler de su stock acumulado de capital físico menos el costo de su inversión corriente.

---

<sup>32</sup> Que se supone constante.

<sup>33</sup> La interpretación empírica de la concavidad del perfil de ganancias con respecto a la experiencia es la depreciación del capital humano ya que las ganancias incrementan hasta cierto punto de la experiencia y luego decrecen si se considera que los individuos en edad adulta son menos productivos.

<sup>34</sup> Para obtener más detalles sobre el tema vease: Willis, Robert (1986), "Wage determinants: A survey and reinterpretation of human capital earning functions" en O. Ashenfelter y R. Layard Handbook of Labor Economics, Elsevier Science, pp. 525-602.



Adicionalmente, Willis (1986) indica que varios estudios que buscan caracterizar la trayectoria del ciclo de vida de las ganancias que se derivan de un programa óptimo de inversión en educación y formación en el puesto de trabajo asumen que el trabajador intenta maximizar el valor presente descontado de los ingresos netos de por vida de los costos directos de inversión. Dicha maximización está sujeta a: i) las limitaciones impuestas por una "función humana de producción de capital", que representa la capacidad del trabajador para transformar los productos de su propio tiempo y los bienes adquiridos en bienes de capital humano y ii) el presupuesto de tiempo a asignar al "aprendizaje" y a la "ganancia". Estos estudios comparten algunas características cualitativas de la función de ganancias de Mincer en (16). En particular, sugieren que el trabajador tiende a especializarse en la inversión en la primera parte de su vida cuando su stock de capital humano es bajo. Esto racionaliza la especialización en educación en el comienzo de la vida. En algún momento, el trabajador combina las ganancias con el aprendizaje y entra en la fuerza de trabajo. En un principio, el trabajador tiende a invertir a una tasa bastante alta porque el nivel de sus ingresos observados es bajo. No obstante, con el paso del tiempo, sus ganancias tienden a crecer rápidamente, tanto por la tasa de acumulación del stock de capital humano y porque el nivel óptimo de inversión disminuye. Eventualmente, la disminución de la tasa de inversión en combinación con la depreciación del stock de capital existente puede resultar una cesación del crecimiento de las ganancias. En este sentido, los ingresos alcanzan un máximo y tienden a disminuir hasta la edad de jubilación.

Así también, para Willis (1986) la justificación **pragmática** de Mincer mencionada consiste en que la función de ganancias del capital humano (16) es una aproximación a los modelos óptimos de capital humano ya que la forma precisa funcional para las ganancias del ciclo de vida implícitas en estos modelos es usualmente desconocida, pues varias características del modelo no son observables tales como el capital humano, el precio de alquiler del capital humano, la tasa de descuento, la forma funcional de la función de producción de capital humano, las entradas de tiempo y bienes comprados utilizados en la inversión, y los parámetros

específicos del individuo de la función de producción que puede ser interpretado como la representación de la interacción de la "capacidad de aprendizaje" del individuo con el hogar, la escuela y los ambientes de trabajo donde el aprendizaje tiene lugar. En concreto, la función de ganancias del capital humano (16) de Mincer (1974) se basa en el supuesto de que por simplicidad la tasa de rendimiento a menudo se trata como un parámetro para el individuo. Esto equivale a suponer que un cambio en la inversión de una persona no cambia su tasa de retorno marginal (promedio).

Para obtener la tasa de crecimiento instantánea de la capacidad de ingresos al tiempo  $t$  Willis (1986) indica que se combinan dos supuestos, el supuesto mencionado en el párrafo anterior con un supuesto sobre la trayectoria temporal de inversión a lo largo del ciclo de vida del individuo para derivar la función de ganancias en (16). En particular, se supone que un individuo comienza con un stock de capital humano de  $E(0)$  a la edad de ingreso a la escuela,  $t=0$ . Se supone también que, en el momento  $t$ , dedica una parte,  $k(t)$ , de su capacidad de generar ingresos a la inversión en capital humano y  $1-k(t)$  a las ganancias, y que  $p$  es la tasa de retorno específica de cada individuo. Por consiguiente, la tasa de crecimiento instantánea de la capacidad de ingresos al tiempo  $t$  de un individuo se expresa así:

$$g(t) = pk(t) \quad (17)$$

Consecuentemente, al tiempo  $t$  la capacidad de ingresos del individuo es:

$$E(t) = E(0)\exp\left\{\int_0^t g(\tau)d\tau\right\} \quad (18)$$

Y los ingresos reales<sup>35</sup> son:

$$y(t) = (1 - k(\tau))E(\tau) \quad (19)$$

Para Willis (1986), la escolaridad es una actividad en la que el individuo se dedica a tiempo completo y su inversión en la misma es  $k(t)=1$  para las edades de  $6$  a  $6+s$ ; es

---

<sup>35</sup>Los ingresos reales son la capacidad de ganancias menos el valor corriente de la inversión.

decir invierte toda su capacidad de generar ingresos en capital humano y no genera ganancias. Consecuentemente, de (19), se deduce que la capacidad de ingresos al salir de la escuela es:

$$E(s) = E(0)e^{ps} \quad (20)$$

Si se considera que el individuo no realiza ninguna inversión adicional después de salir de la escuela; es decir,  $k(t)=0$  para  $t > s$ , el perfil de ganancias individual del ciclo de vida sería horizontal con un valor de  $y(s)=E(s)$ . Por consiguiente, calculando el logaritmo neperiano de ambos lados de la ecuación (20), se obtiene la relación escolaridad-ganancias que es de la forma log-lineal y es la función minceriana bajo el supuesto de que la tasa de retorno es constante para todos los individuos; en consecuencia la función mencionada es de la siguiente manera:

$$\ln[E(s)] = \ln[E(0)e^{ps}]$$

$$\ln[E(s)] = ps + \ln[E(0)]$$

$$\ln y = ps + \ln[E(0)] \quad (21)$$

$$\ln y_s = \ln y_0 + ps \quad (22)$$

Donde,  $Y_s$  es el nivel de ingresos de las personas que no invierten en capital humano más allá de  $s$  años de escolaridad y  $Y_0$  los ingresos de los individuos sin un nivel de educación.

Según Willis (1986), las teorías de la acumulación óptima de capital humano sugieren que el monto de la inversión que los trabajadores destinan en la formación en el puesto de trabajo después de salir de la escuela, disminuye con el tiempo. Por consiguiente, partiendo de la ecuación (16) que corresponde (aproximadamente) a la suposición de que la fracción de la capacidad de ingresos que se invierte declina linealmente durante la vida de trabajo desde un valor inicial de  $k(0)$  al comienzo de la carrera de trabajo a un valor de cero al final de la carrera y sea  $k(x)=k(0)-(k(0)/n) x$ ,

donde  $n$  es la longitud de la vida de trabajo. En este caso, la capacidad de ingresos es:

$$E(x) = E(s) \exp\left\{p \int_0^x [k(0) - (k(0)/n)t] dt\right\}$$

$$E(x) = E(s) \exp\{pk(0)x - (pk(0)/2n)x^2\} \quad (23)$$

Los ingresos actuales netos de los costos de inversión son  $y(x) = (1-k(x)) E(x)$ . Así, reemplazando (21) y (23) implican que:

$$\ln y = \ln(1 - k(x)) + \ln(E(x))$$

$$\ln y = \ln(1 - k(x)) + \ln(E(s)) + pk(0)x - \left(\frac{pk(0)}{2n}\right)x^2$$

$$\ln y = \ln(E(0)) + ps + pk(0)x - (pk(0)/2n)x^2 + \ln(1 - k(x)) \quad (24)$$

En conclusión, Mincer considera a la función (16) como una aproximación de la ecuación (24).

## 3.2. PRINCIPALES PROBLEMAS ECONOMÉTRICOS EN LA ESTIMACIÓN DE ECUACIONES MINCERIANAS

### 3.2.1 PROBLEMAS EMPÍRICOS EN LA ESTIMACIÓN DE ECUACIONES MINCERIANAS

Las ecuaciones mincerianas se basan en ciertos supuestos, por lo que al momento de las estimaciones existen diferentes problemas estadísticos y econométricos; a continuación se describen algunos de ellos.

Willis (1986) indica que uno de los supuestos que sustenta el uso de una función estadística de los ingresos para estimar la tasa de retorno a la educación es que

representa con precisión el conjunto de posibilidades que enfrenta un individuo (después de controlar por características observables exógenas tales como la raza o el sexo); pero la literatura sobre la inversión en capital humano se preocupa de que la función estadística de ganancias no mida correctamente el conjunto de posibilidades individuales. Por consiguiente, la tasa estimada de retorno de la educación puede ser altamente sesgada, entre otros argumentos, por la capacidad individual que es inobservable provocando que los individuos con "alta capacidad", en promedio, tengan un mayor nivel de escolaridad que los individuos con "baja capacidad". Si es así, la  $u$  residual, en la ecuación (6) se correlaciona positivamente con los años de escolaridad,  $s$ , y la función de rentabilidad estimada estará sujeta a un "sesgo de habilidad" que sobrestima las ganancias que lograría una persona con capacidades a través del incremento de la escolarización.

Por otra parte, Willis (1986) identifica otros problemas fundamentales que son: (a) es imposible observar los beneficios del ciclo de vida de la misma persona que ha hecho inversiones en educación alternativa (o después de la escuela) y (b) es imposible observar todas las variables (por ejemplo, las habilidades) que determinan las posibilidades de ingresos. A lo sumo, se observa la trayectoria de ganancias de un individuo que ha elegido (o asignado) un determinado nivel de escolaridad. Por lo tanto, cualquier medida de la rentabilidad de la inversión debe basarse en la comparación de los ingresos de los diferentes individuos que difieren en los niveles de escolaridad.

Pero para Willis (1986) si los niveles de escolaridad (y las inversiones después de la escuela) hubieran sido asignados al azar para cada grupo de capacidad de acuerdo a un diseño experimental, la función estadística de ganancias estimada a partir de las diferencias interpersonales en los ingresos, la educación y la experiencia proporcionaría una estimación no sesgada del conjunto de posibilidades de un

individuo en ese grupo<sup>36</sup>, ya que, por diseño, el término de error,  $u$ , sería independiente de los años de escolarización ( $s$ ) y la experiencia ( $x$ ).

Sin embargo, Willis (1986) menciona adicionalmente, que otro problema para la estimación del modelo de escolaridad es la hipótesis básica del comportamiento de la economía, la cual consiste en que los agentes económicos seleccionan la alternativa preferida de su conjunto de posibilidades. Pero, si el conjunto de todas las posibilidades que no pueden ser observadas y si las posibilidades varían a través de agentes, el acto de elección óptima implica que los datos del mercado son sistemáticamente censurados y no hay ninguna garantía de que las estimaciones basadas en las diferencias interpersonales en los ingresos y escolaridad estimaran con precisión el conjunto de posibilidades de cualquier individuo de la población. En el contexto de la literatura sobre la inversión en la educación, esto ha llegado a ser conocido como el "**problema de autoselección**". Este problema es considerado como omnipresente en la economía y presenta difíciles problemas econométricos en cualquier situación en la que el conjunto de todas las oportunidades de cada agente no se observe. Asimismo, muchas de las cuestiones empíricas (incluida la cuestión de sesgo de habilidad) que han surgido en la literatura de la función de los ingresos pueden ser interpretadas en términos del problema de auto-selección.

Otro de los problemas que Harmon et al. (2003) hacen mención es que la medida de escolarización es tratada como exógena y en la teoría del capital humano subyacente, la educación es una variable de elección endógena. Es útil, por tanto, considerar las implicaciones de la endogeneidad de la escolaridad.

Además, Harmon et al. (2003) manifiestan que en el modelo básico, la tasa interna de retorno de la educación está determinada en parte por los ingresos no percibidos (menos cualquier subsidio del gobierno o contribuciones de los padres) y los gastos de educación. Al presentar diferencias de capacidad tiene dos efectos en este cálculo básico. Los individuos más capaces pueden "convertir" la escolarización en

---

<sup>36</sup>Por ejemplo, proporcionaría la mejor estimación de la diferencia en los ingresos del ciclo de vida que un individuo podría esperar dados los niveles alternativos de educación.

capital humano de manera más eficiente que los menos capaces, y esto plantea la TIR para los más capaces. Se podría pensar en esto como una habilidad inherente y la educación como factores complementarios en la producción de capital humano a fin de que, para un incremento dado en educación, una mayor dotación de capacidad genera más capital humano. Por otro lado, los más capaces pueden tener mayores costos de oportunidad ya que pueden haber sido aptos para ganar más en el mercado de trabajo, si la capacidad de progresar en la escuela se correlaciona positivamente con la capacidad de ganar, y esto reduce la TIR.

Por otra parte, empíricamente para Harmon et al. (2003), el sesgo potencial en las estimaciones de mínimos cuadrados del rendimiento de la educación a ser derivada de la ecuación minceriana estará determinado por la varianza en la capacidad relativa (habilidad), por la varianza en las tasas de descuento, así como también por la covarianza entre ellos. Este sesgo de "endogeneidad" se debe a que las personas con mayores retornos marginales o menores costos marginales de la educación elijen mayores niveles de escolaridad. Si no hay variación en la tasa de descuento, luego la endogeneidad surgirá únicamente a partir de la correlación entre la capacidad con la educación y dado que es probable que la correlación sea positiva, el sesgo en las estimaciones de mínimos cuadrados ordinarios (MCO) será hacia arriba (si aumenta la capacidad de los salarios en el futuro más de lo que aumenta los salarios iniciales). Si no hay variación en la capacidad, el carácter endógeno surge únicamente de la correlación (negativa) entre las tasas de descuento y la cantidad de la educación, por lo que las estimaciones por mínimos cuadrados ordinarios (MCO) serán sesgadas hacia abajo si las tasas de descuento y los salarios están positivamente correlacionados (por ejemplo, si la gente ambiciosa gana salarios más altos y son más impacientes). Por lo tanto, la dirección del sesgo en las estimaciones MCO de los rendimientos de la educación no es clara y es, en última instancia, una cuestión empírica.

Adicionalmente, Harmon et al. (2003) indican que en el modelo de Mincer (1974), el término de error captura los efectos individuales inobservables y estos factores

individuales también pueden influir en la decisión de la escolaridad, y por lo tanto inducir una correlación entre la educación y el término de error en la función de los ingresos. Un ejemplo común es la habilidad no observada. Por consiguiente si la escolaridad es endógena, los resultados obtenidos por el método de mínimos cuadrados ordinarios proporcionan estimaciones sesgadas del rendimiento de la educación.

Otro problema en relación a la endogeneidad de la educación que Harmon et al. (2003) reconocen es que la educación puede tener un valor en el mercado de trabajo no debido a algún efecto en la productividad, sino de falsas razones. En particular, la educación puede actuar como una señal de capacidad (u otras características que los empleadores valoran, ya que contribuye a la productividad, pero que no pueden observar fácilmente). Por simplicidad supóngase que los empleadores creen que la educación está relacionada con la productividad, entonces esto va a ser confirmado por su experiencia si se da el caso de que los individuos de alta productividad eligen altos niveles de educación. Esto será cierto si los costos de educación son lo suficientemente bajos para los individuos de alta productividad, mientras que para los individuos de baja productividad estos costos serían mayores. Así, el mercado se caracteriza por un equilibrio separador donde los individuos de alta productividad eligen altos niveles de educación en estas condiciones específicas. Además, es posible que el rendimiento de la educación refleje que la educación es una señal de la productividad inherente de la persona y no un medio para mejorar la productividad.

Analíticamente, Harmon et al. (2003) explican el problema de endogeneidad mediante las ecuaciones de ingresos de la siguiente manera:

$$\log y_i = X_i\beta + rs_i + u_i \quad (25)$$

$$s_i = Z_i\alpha + v_i \quad (26)$$

Donde  $y_i$  es el ingreso individual,  $X_i$  y  $Z_i$  son vectores de características del individuo  $i$ ,  $s_i$  corresponde a los años de escolaridad,  $r$  define el efecto causal de la escolaridad y  $u_i$ ,  $v_i$  son los términos de error.



Según Harmon et al. (2003), la estimación de la ecuación de salarios (25) por MCO permite obtener un  $\beta$  no sesgado, sólo si  $s_i$  es exógena, por lo tanto, no hay correlación entre los dos términos de error  $u_i$  y  $v_i$ . Si esta condición no se cumple las estimaciones por MCO estarán sesgadas. Particularmente, la correlación puede ser distinta de cero debido a que algunas variables importantes relacionadas con la educación y los ingresos son omitidas del vector  $X$ . Entre otras variables, las comúnmente desplazadas o no consideradas son la motivación o las medidas de capacidad (como el coeficiente intelectual ( $CI$ )). Es importante señalar que incluso una lista muy amplia de variables incluidas en el vector  $X$  nunca serán medidas exhaustivas de las características de los individuos. En este sentido, las estimaciones del rendimiento de la educación basada en MCO no darán el efecto causal de la educación sobre las ganancias ya que el coeficiente de escolarización  $\beta$  captura algunos de los efectos que de otra manera serían atribuidos a las variables omitidas. Por ejemplo, si la variable omitida es la motivación, y si tanto la escolaridad y los ingresos se correlacionan positivamente con la motivación, la estimación de MCO ignora que las personas más motivadas tienden a ganar más que personas menos motivadas incluso cuando tienen cantidades similares de escolaridad.

Otro problema econométrico según Harmon et al. (2003) es que en la economía existe una gran heterogeneidad en los rendimientos de la educación pues los rendimientos pueden diferir entre individuos, ya que difieren en la eficiencia con la que pueden explotar la educación para elevar su productividad. Además, los rendimientos de la educación pueden ser diferentes en toda la distribución salarial.

Por otra parte, Heckman (1979) desde una perspectiva econométrica, detalla el problema de selección muestral que se muestra a continuación. Primeramente considera dos modelos de ecuaciones y una muestra aleatoria de  $I$  individuos. Las ecuaciones de ingresos correspondientes al individuo  $i$  son:

$$Y_{1i} = X_{1i}\beta_1 + U_{1i} \tag{27a}$$

$$Y_{2i} = X_{2i}\beta_2 + U_{2i} \quad (27b)$$

Para  $i=1,2,\dots,l$ .

Donde:

$X_{ji}$ : es un vector ( $1 \times K_j$ ) de variables exógenas; y

$\beta_j$ : un vector de parámetros ( $K_j \times 1$ ).

Además las ecuaciones (27a) y (27b) se establecen bajo los supuestos:

$$E(U_{ji}) = 0,$$

$$E(U_{ji}, U_{j'i'}) = \sigma_{jj'}, \text{ para } i=i'; \text{ y}$$

$$E(U_{ji}, U_{j'i'}) = 0, \text{ para } i \neq i'.$$

Adicionalmente, para Heckman (1979) la densidad conjunta de  $U_{1i}, U_{2i}$  es  $h(U_{1i}, U_{2i})$ . También, la matriz de regresores es de rango completo<sup>37</sup> de manera que si todos los datos estaban disponibles, los parámetros de cada ecuación pueden ser estimados por el método de mínimos cuadrados ordinarios.

Por otra parte, Heckman (1979) supone que se busca estimar la ecuación (27a) pero esos datos están perdidos sobre  $Y_1$  para ciertas observaciones. Por consiguiente, la función de regresión de la población para la ecuación (27a) se puede escribir como:

$$E(Y_{1i} | X_{1i}) = X_{1i}\beta_1, \text{ para } i=1,2,\dots,l.$$

La función de regresión para la submuestra de los datos disponibles es:

$$\begin{aligned} E(Y_{1i} | X_{1i}, \text{regladeselección delamuestra}) \\ = X_{1i}\beta_1 + E(U_{1i} | \text{regladeselección delamuestra}) \end{aligned}$$

---

<sup>37</sup> Una matriz es de **rango completo** cuando todas las columnas son independientes, es decir, las columnas no se combinan linealmente entre ellas. Para una ampliación del tema véase Castro (2008), pp. 189-192.

para  $i=1,2,\dots,l$ , donde se adopta la convención de que el primero  $l_1 < l$  observaciones tienen datos disponibles en  $Y_{1i}$ .

Heckman (1979) también indica que si la esperanza condicional de  $U_{1i}$  es cero ( $E(U_{1i}|\text{regladeselección delamuestra}) = 0$ ), la función de regresión para la submuestra seleccionada es la misma que la función de regresión de la población ( $Y_{1i} = E(Y_{1i}|X_{1i}, \text{regladeselección delamuestra}) = X_{1i}\beta_1$ ). Además, los estimadores de mínimos cuadrados pueden ser utilizados para estimar  $\beta_1$  en la submuestra seleccionada. Por lo demás, el único coste de tener una muestra incompleta es una pérdida de eficiencia.

Adicionalmente, Heckman (1979) reconoce que en el caso general, la regla de selección de muestra que determina la disponibilidad de datos trae consecuencias negativas. Así por ejemplo, suponiendo que se dispone de datos sobre  $Y_{1i}$  si  $Y_{2i} \geq 0$ , mientras que si  $Y_{2i} < 0$ , no hay observaciones sobre  $Y_{1i}$ <sup>38</sup>. Analíticamente, la representación del caso general es la siguiente:

$$\begin{aligned} E(U_{1i}|X_{1i}, \text{regladeselección delamuestra}) &= E(U_{1i}|X_{1i}, Y_{2i} \geq 0) \\ &= E(U_{1i}|X_{1i}, U_{2i} \geq -X_{2i}\beta_2) \end{aligned}$$

Para Heckman (1979), en el caso de independencia entre  $U_{1i}$  y  $U_{2i}$ , de tal forma que los datos sobre  $Y_{1i}$  están aleatoriamente perdidos, la media condicional de  $U_{1i}$  es cero. En el caso general, es distinto de cero y la función de regresión de la submuestra es:

$$E(Y_{1i}|X_{1i}, Y_{2i} \geq 0) = X_{1i}\beta_1 + E(U_{1i}|U_{2i} \geq -X_{2i}\beta_2) \quad (28)$$

Heckman (1979) además observa que uno de los problemas econométricos que hace referencia a que los parámetros de la ecuación (27a) obtenidos de la muestra seleccionada omiten el término final de la ecuación (28) como regresor, de modo que el sesgo que resulta de usar las muestras seleccionadas de manera no aleatoria para

---

<sup>38</sup> La elección de cero como un umbral implica una normalización no esencial.

estimar las relaciones de comportamiento, surge del problema ordinario de variables omitidas.

Adicionalmente, según Heckman (1979) el sesgo de selección muestral puede surgir por dos razones. En primer lugar, puede existir **autoselección** por los individuos o por las unidades de datos que están siendo investigadas. En segundo lugar, las decisiones de selección de la muestra de los analistas o los procesadores de datos funcionan de la misma manera que la autoselección.

### 3.2.2 CORRECCIÓN DE LOS PRINCIPALES PROBLEMAS ECONOMETRÍCOS EN LA ESTIMACIÓN DE ECUACIONES MINCERIANAS

Luego de la expresión analítica y la descripción de los principales problemas econométricos y empíricos en las estimaciones de ecuaciones mincerianas; en esta sección se detalla la solución a los problemas mencionados según Heckman (1979). Asumiendo que  $h(U_{1i}, U_{2i})$  es una densidad normal bivariada y utilizando los siguientes resultados conocidos:

$$E(U_{1i}|U_{2i} \geq -X_{2i}\beta_2) = \frac{\sigma_{12}}{(\sigma_{22})^{1/2}} \lambda_i$$

$$E(U_{2i}|U_{2i} \geq -X_{2i}\beta_2) = \frac{\sigma_{22}}{(\sigma_{22})^{1/2}} \lambda_i$$

Donde:

$$\lambda_i = \frac{\phi(Z_i)}{1 - \Phi(Z_i)} = \frac{\phi(Z_i)}{\Phi(-Z_i)}$$

Sea  $\phi$  la función de densidad y  $\Phi$  la función de distribución para una variable normal estándar; también se tiene a  $Z_i$  que representa la siguiente igualdad:

$$Z_i = -\frac{X_{2i}\beta_2}{(\sigma_{22})^{1/2}}$$

Heckman (1979) indica que “ $\lambda_i$ ” es la inversa del ratio de Mills y es una función monótona decreciente de la probabilidad de que una observación es seleccionada de la muestra  $\Phi(-Z_i)$  ( $=1- \Phi(Z_i)$ ).

Seguidamente, se escribe la función de regresión condicional para muestras elegidas de la siguiente manera:

$$E(Y_{1i}|X_{1i}, Y_{2i} \geq 0) = X_{1i}\beta_1 + \frac{\sigma_{12}}{(\sigma_{22})^{1/2}} \lambda_i$$

$$E(Y_{2i}|X_{2i}, Y_{2i} \geq 0) = X_{2i}\beta_2 + \frac{\sigma_{22}}{(\sigma_{22})^{1/2}} \lambda_i$$

$$Y_{1i} = E(Y_{1i}|X_{1i}, Y_{2i} \geq 0) + V_{1i} \quad (29a)$$

$$Y_{2i} = E(Y_{2i}|X_{2i}, Y_{2i} \geq 0) + V_{2i} \quad (29b)$$

Bajo los siguientes supuestos:

$$E(V_{1i}|X_{1i}, \lambda_i, U_{2i} \geq -X_{2i}\beta_2) = 0 \quad (29c)$$

$$E(V_{2i}|X_{2i}, \lambda_i, U_{2i} \geq -X_{2i}\beta_2) = 0 \quad (29d)$$

$$E(V_{ji}V_{j'i'}|X_{1i}, X_{2i}, \lambda_i, U_{2i} \geq -X_{2i}\beta_2) = 0, \text{ para } i \neq i' \quad (29e)$$

$$E(V_{1i}^2|X_{1i}, \lambda_i, U_{2i} \geq -X_{2i}\beta_2) = \sigma_{11} \left( (1 - \rho^2) + \rho^2(1 + Z_i\lambda_i - \lambda_i^2) \right) \quad (29f)$$

$$E(V_{1i}V_{2i}|X_{1i}, X_{2i}, \lambda_i, U_{2i} \geq -X_{2i}\beta_2) = \sigma_{12}(1 + Z_i\lambda_i - \lambda_i^2) \quad (29g)$$

$$E(V_{2i}^2|X_{2i}, \lambda_i, U_{2i} \geq -X_{2i}\beta_2) = \sigma_{22}(1 + Z_i\lambda_i - \lambda_i^2) \quad (29h)$$

Donde,

$$\rho^2 = \frac{\sigma_{12}^2}{\sigma_{11}\sigma_{22}}$$

$$0 \leq 1 + Z_i\lambda_i - \lambda_i^2 \leq 1 \quad (30)$$

Si se conoce  $Z_i$  y por consiguiente  $\lambda_i$ , Heckman (1979) sugiere que  $\lambda_i$  se introduzca como un regresor en la ecuación (29a) y luego estimar dicha ecuación por el método de mínimos cuadrados ordinarios. El estimador de mínimos cuadrados de  $\beta_1$  y  $\sigma_{12}/(\sigma_{22})^{1/2}$  es no sesgado pero es ineficiente. La ineficiencia es una consecuencia de la heterocedasticidad evidente a partir de la ecuación (29f) cuando  $X_{2i}$  (y por lo tanto  $Z_i$ ) contiene regresores no triviales. Además, como consecuencia de la desigualdad (30), el estimador estándar de mínimos cuadrados de la varianza de la población  $\sigma_{11}$  es sesgado hacia abajo. También, como consecuencia de la ecuación (29g) y la desigualdad (30), el estimador usual de la covarianza está sesgado hacia abajo.

Es importante recalcar que en la práctica no se conoce  $\lambda_i$ . Pero si existe el caso de que la muestra está censurada, en la cual no hay información sobre  $Y_{1i}$  si  $Y_{2i} \leq 0$ , pero si se conoce  $X_{2i}$  para las observaciones con  $Y_{2i} \leq 0$ , se podría estimar  $\lambda_i$  mediante el siguiente procedimiento que Heckman (1979) plantea:

- i. Estimar los parámetros de la probabilidad de que  $Y_{2i} \geq 0$  (es decir,  $\beta_2/(\sigma_{22})^{1/2}$ ) utilizando el análisis probit para la muestra completa.
- ii. Después de estimar  $\beta_2^*$  que es igual a  $\beta_2/(\sigma_{22})^{1/2}$ ; se puede estimar  $Z_i$  y por ende  $\lambda_i$ . Estos estimadores son consistentes<sup>39</sup>.
- iii. El valor estimado de  $\lambda_i$  puede ser usado como un regresor en la ecuación (29a). Los estimadores de regresión de la ecuación (29a) son consistentes para  $\beta_1$  y  $\sigma_{12}/(\sigma_{22})^{1/2}$  (los coeficientes de  $X_{1i}$  y  $\lambda_i$  respectivamente).<sup>40</sup>
- iv. Se puede estimar consistentemente  $\sigma_{11}$  mediante el siguiente procedimiento. A partir del paso iii., estimar de forma consistente  $C = \rho(\sigma_{11})^{1/2} = \sigma_{12}/(\sigma_{22})^{1/2}$ . Siendo  $\hat{V}_{1i}$  la notación del residuo para la observación  $i$ -ésima obtenido en la etapa iii. y  $\hat{C}$  del estimador de  $C$ . Así también, el estimador de  $\sigma_{11}$  es:

<sup>39</sup>Se dice que un estimador es **Consistente** cuando a medida que el tamaño de la muestra crece, el valor del estimador tiende a ser el valor del parámetro.

<sup>40</sup>Asumiendo que el vector  $X_{2i}$  contiene regresores no triviales o que  $\beta_1$  no contiene ninguna intersección o ambos.

$$\hat{\sigma}_{11} = \frac{\sum_{i=1}^{I_1} \hat{V}_{1i}^2}{I_1} - \frac{\hat{C}}{I_1} \sum_{i=1}^{I_1} (\hat{\lambda}_1 \hat{Z}_i - \hat{\lambda}_i^2)$$

Donde  $\hat{\lambda}_i$  y  $\hat{Z}_i$  son los valores estimados de  $\lambda_i$  y  $Z_i$  respectivamente, estos valores se obtienen del paso ii. Adicionalmente, Heckman (1979) destaca que el estimador  $\sigma_{11}$  es consistente y positivo ya que el término de la segunda sumatoria debe ser negativo (véase la desigualdad (30)).

Heckman (1979) explica que las fórmulas habituales para los errores estándar de los coeficientes obtenidos por el método de mínimos cuadrados no son apropiados, excepto en el caso de la hipótesis nula de ausencia de sesgo de selección ( $C = \sigma_{12}/(\sigma_{22})^{1/2} = 0$ ). En ese caso, los errores de la regresión estándar usual son apropiados y una prueba exacta de la hipótesis nula  $C = 0$  se puede realizar usando la distribución t. Si  $C \neq 0$ , el procedimiento usual para calcular los errores estándar subestima los verdaderos errores estándar y exagera los niveles de significancia estimados<sup>41</sup>.

### 3.3. METODOLOGÍAS PARA LA DEFINICIÓN EMPÍRICA DE LOS FACTORES OBSERVABLES Y NO OBSERVABLES DE LOS DIFERENCIALES SALARIALES POR SECTOR

En el presente apartado, se describe la metodología a utilizarse, la cual se basa en tres procedimientos ya mencionados en el trabajo de investigación de Torresano (2009); (i) la estimación de ecuaciones mincerianas de ingresos controladas por

---

<sup>41</sup> Además, Heckman (1979) indica que la ecuación (29a) utiliza un valor estimado de  $\lambda_i$  en lugar del valor real de  $\lambda_i$ ; por tanto, tomando en cuenta que el término de error en la ecuación son los dos últimos términos de la misma, esta se representa de la siguiente manera:  $Y_{1i} = X_{1i}\beta_1 + C\hat{\lambda}_i + C(\lambda_i - \hat{\lambda}_i) + V_{1i}$

sector público y privado y su interpretación mediante el análisis de sus coeficientes<sup>42</sup>; (ii) estimación de un modelo multiplicativo de ingresos, igualmente controlado por sector público y privado, con interacciones para que los coeficientes varíen por sector; y (iii) en el último proceso se aplicará el método de Blinder – Oaxaca (1973), utilizando una descomposición de ingresos. Conjuntamente en el transcurso de la investigación se aplicará un modelo econométrico que permita la corrección de sesgo de selección muestral; por lo tanto se utilizará el modelo planteado por Heckman (1979).

### 3.3.1 MODELOS ADITIVOS

El modelo aditivo controlado por sector económico que se aplicará es una extensión de la ecuación (16), el cual muestra que las diferencias salariales son únicamente en nivel, pues de inicio hay una diferencia que se mantiene en el tiempo. Así, el modelo a aplicarse es de la siguiente forma:

$$W_i = \ln(Y_i) = \beta X_i + \alpha P_i + \mu_i \quad (31)$$

Donde:

$W_i$  = logaritmo del ingreso por hora del individuo  $i$ , el cual es calculado por el logaritmo natural de la división de los ingresos mensuales laborales corrientes, tanto de actividades principales y secundarias y las horas mensuales trabajadas por empleados dependientes e independientes<sup>43</sup>.

---

<sup>42</sup> Las ecuaciones mincerianas ya son expuestas en el apartado 3.1.

<sup>43</sup> El análisis de los salarios se realiza para un solo instante de tiempo, específicamente para diciembre de 2011, ya que se utilizará únicamente la Encuesta de Empleo y Desempleo de dicha fecha. Es importante mencionar, que muchas de las aplicaciones de la teoría de capital humano y análisis discriminatorio utilizan generalmente análisis de corte transversal dada la disponibilidad de información y por el potencial de sus resultados. Sin embargo, es factible realizar análisis discriminatorios para diversos espacios de tiempo pero esta investigación se limita a un solo período.



$X_i$ = matriz de características observables del individuo  $i$ , tales como años de educación, experiencia potencial<sup>44</sup> (y experiencia potencial al cuadrado, tal como se detalla en la ecuación (16), género, etnia, relación de parentesco con el jefe de hogar, situación migratoria y localización geográfica.

$P_i$ = variable indicatriz que toma el valor de 1 si el individuo  $i$  pertenece al sector público y cero si pertenece al sector privado, y

$\mu_i$ = término de error.

El presente modelo puede ser estimado mediante Mínimos Cuadrados Ordinarios (MCO) y si en sus resultados  $\alpha < 0$ , entonces se sospecharía de la existencia de discriminación laboral en contra de los trabajadores del sector público.

Este modelo, debe ser corregido por posibles problemas sesgo muestral<sup>45</sup> con el modelo Heckman (1979) que incorpora “la razón de Mills” como regresor adicional, así:

$$W_i = \ln(Y_i) = \beta X_i + \alpha P_i + \hat{\lambda} \sigma + n_i \quad (32)$$

Donde:

$W_i, X_i, P_i$  son términos ya definidos en la ecuación (31),

$\sigma$ = covarianza entre el termino de error de la ecuación de ingresos (32) y participación laboral (33),

$n_i$ = término residual, y

---

<sup>44</sup> Según Altonji y Blank (1999), la experiencia potencial para cada individuo puede ser calculada como el resultado de la diferencia de la edad - los años de educación - 5 y bajo el supuesto de que las personas están trabajando en su edad adulta cuando no están en la escuela.

<sup>45</sup> El sesgo de selección muestral puede ser consecuencia de que en la muestra solo se consideren a un grupo de individuos con características específicas como por ejemplo se pueden excluir de la muestra a los trabajadores que no participan en el mercado laboral.

$\hat{\lambda}_i$  = inversa de la razón de Mills que proviene de la estimación del modelo probit de participación laboral que se expresa de la siguiente manera:

$$\pi_i = \Phi(\beta X_i) \quad (33)$$

Donde,

$\pi_i$  = probabilidad de estar ocupado,

$\Phi$  = función de distribución normal con parámetros  $N(0, 1)$ ; además,

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp\left(-\frac{1}{2}t^2\right) dt$$

$\beta$  = vector fila de parámetros desconocidos, y

$X_i$  = matriz de variables observables individuales (como educación, experiencia, género, etnia, número de miembros del hogar, estado civil, parentesco y localización geográfica).

Además, Torresano (2009) indica que si  $\hat{\lambda}$  es cero, entonces el término de ajuste será cero y por ende el MCO es consistente e insesgado caso contrario el estimador del modelo de selección muestral es consistente pero no eficiente dado que el componente de error  $\eta_i$  es heterosedástico.

### 3.3.2 MODELOS MULTIPLICATIVOS

El siguiente modelo que se detalla a continuación es el denominado modelo multiplicativo de ingresos controlado por sector público y privado, en el cual según Torresano (2009) se introducen interacciones con algunas variables de control y su finalidad es que los coeficientes varíen por sector. Así el modelo se presenta de la siguiente manera:

$$W_i = \beta X_i + \alpha P_i + \gamma X_i P_i + \mu_i \quad (34)$$

Donde:

$W_i$ = logaritmo del ingreso por hora del individuo  $i$ ,

$X_i$ = matriz de características observables del individuo  $i$  (similares a las descritas en el punto anterior),

$P_i$ = variable dicotómica que toma el valor de 1 si el individuo  $i$  pertenece al sector público y cero si pertenece al sector privado,

$\mu_i$ = término de error, y los parámetros pueden ser interpretados así:

$\alpha$ = diferencia inicial de salarios y si  $\alpha > 0$ , entonces existiría un premio a favor del sector privado; caso contrario existiría una penalización.

$\gamma$ = pendientes de las rectas.

El modelo multiplicativo al igual que el aditivo puede ser estimado mediante Mínimos Cuadrados Ordinarios y luego de la estimación se realiza la corrección de sesgo muestral, utilizando el modelo econométrico planteado por Heckman (1979), así:

$$W_i = \beta X_i + \alpha P_i + \gamma X_i P_i + \hat{\lambda} \sigma + n_i \quad (35)$$

Donde:

$W_i, X_i, P_i$  son términos ya identificados en la ecuación (31).

$\hat{\lambda}_i$ = inversa de la razón de Mills.

$\sigma$ = covarianza entre el término de error de la ecuación de ingresos y participación laboral.

$n_i$ = término residual.

### 3.3.3 DESCOMPOSICIÓN DE INGRESOS DE BLINDER-OAXACA

Adicionalmente a los dos procedimientos anteriores, para analizar el fenómeno en estudio, también se utilizará la denominada Descomposición de Ingresos de Blinder-

Oaxaca (1973); este método puede ser estimado por Mínimos Cuadrados Ordinarios y es corregido por el problema de sesgo de selección muestral a través del modelo planteado por Heckman (1979).

Para Blinder (1972), el análisis de la dispersión de los salarios individuales inicia con el planteamiento de la siguiente ecuación:

$$\ln(Y_i) = \beta_0 + \sum_{j=1}^n \beta_j X_{ji} + \mu_i \quad (36)$$

para toda la muestra, donde:

$Y_i$  = ingresos o tasa salarial,

$X_{1i}, \dots, X_{ni}$  =  $n$  características observables usadas para explicar  $Y$  (similares a las descritas en los puntos anteriores, pero diferenciadas por grupo), y

$\mu_i$  = término de error.

Dado que la descomposición de ingresos que se desea observar es de dos grupos: sector público y privado, la ecuación (36) se descompone en 2 ecuaciones como son:

$$\ln(Y_i^{pr}) = \beta_0^{pr} + \sum_{j=1}^n \beta_j^{pr} X_{ji}^{pr} + \mu_i^{pr} \quad (36a)$$

$$\ln(Y_i^{pl}) = \beta_0^{pl} + \sum_{j=1}^n \beta_j^{pl} X_{ji}^{pl} + \mu_i^{pl} \quad (36b)$$

Donde el superíndice  $pl$  indica el grupo que gana altos salarios o grupo mayoritario, en el presente estudio hace referencia al Sector Público, y el superíndice  $pr$  indica el grupo que gana bajos salarios o grupo minoritario, el cual en la investigación representa al Sector Privado.

Al igual que los casos anteriores, debido a la existencia de un posible problema de sesgo de selección muestral y para corregir dicho problema a las ecuaciones (36a) y (36b) se añade el ratio de Mills así:

$$\ln(Y_i^{pr}) = \beta_0^{pr} + \sum_{j=1}^n \beta_j^{pr} X_{ji}^{pr} + \widehat{\lambda}^{pr} \sigma^{pr} + \mu_i^{pr} \quad (36c)$$

$$\ln(Y_i^{pl}) = \beta_0^{pl} + \sum_{j=1}^n \beta_j^{pl} X_{ji}^{pl} + \widehat{\lambda}^{pl} \sigma^{pl} + \mu_i^{pl} \quad (36d)$$

Donde los términos de las ecuaciones (36c) y (36d) son conocidos.

Luego, obteniendo la diferencia entre (36a) y (36b) se procede a la descomposición de las diferencias salariales entre el sector público y privado considerando como estructura salarial no discriminatoria a la del sector público; dicha descomposición de ingresos es de la siguiente manera:

$$\sum_j \beta_j^{pr} \bar{X}_j^{pr} - \sum_j \beta_j^{pl} \bar{X}_j^{pl} = \sum_j \beta_j^{pl} (\bar{X}_j^{pr} - \bar{X}_j^{pl}) + \sum_j \bar{X}_j^{pr} (\beta_j^{pr} - \beta_j^{pl}) \quad (37)$$

Donde  $\bar{X}_j$  hace referencia a la característica media de los individuos.

Para la interpretación, Blinder (1972) descompone la parte derecha de la ecuación (37) en dos partes; siendo  $E = \sum_j \beta_j^{pl} (\bar{X}_j^{pr} - \bar{X}_j^{pl})$  el primer elemento y  $C = \sum_j \bar{X}_j^{pr} (\beta_j^{pr} - \beta_j^{pl})$  el segundo. Por lo tanto  $E$  es la porción del diferencial atribuible a las dotaciones diferentes o es considerado según Torresano (2009) como la parte de las diferencias de ingresos explicadas por las diferencias en las dotaciones grupales; por otra parte,  $C$  es la porción del diferencial atribuible a los diferentes coeficientes; de igual forma Torresano (2009) lo interpreta como la parte que mide la contribución de las diferencias en los coeficientes (incluido las diferencias en los interceptos) asociado con el efecto de la discriminación laboral en contra del grupo minoritario que en este caso corresponde al sector público.

Adicionalmente, la descomposición de las diferencias salariales entre el sector público y privado considerando como estructura salarial no discriminatoria a la del sector privado se muestra a continuación:

$$\sum_j \beta_j^{pr} \bar{X}_j^{pr} - \sum_j \beta_j^{pl} \bar{X}_j^{pl} = \sum_j \beta_j^{pr} (\bar{X}_j^{pr} - \bar{X}_j^{pl}) + \sum_j \bar{X}_j^{pl} (\beta_j^{pr} - \beta_j^{pl}) \quad (38)$$

Esta última ecuación (38) se descompone e interpreta de forma similar a la ecuación (37).

Dado que una parte de la brecha salarial entre los sectores económicos está vinculada con el proceso de selección, López et al. (2008) sugieren que para la corrección de dicho problema de sesgo de selección se añade a la ecuación (37)<sup>46</sup> (que considera como estructura salarial no discriminatoria a la del sector público) el ratio de Mills, así:

$$\sum_j \beta_j^{pr} \bar{X}_j^{pr} - \sum_j \beta_j^{pl} \bar{X}_j^{pl} = \sum_j \beta_j^{pl} (\bar{X}_j^{pr} - \bar{X}_j^{pl}) + \sum_j \bar{X}_j^{pr} (\beta_j^{pr} - \beta_j^{pl}) + (\bar{\lambda}^{pr} \sigma^{pr} - \bar{\lambda}^{pl} \sigma^{pl}) \quad (37a)$$

Donde la barra sobre las variables denota su media y  $\bar{\lambda}^{pl}$  y  $\bar{\lambda}^{pr}$  es la media de la inversa de la ratio de Mills para los individuos que pertenecen al sector público y privado respectivamente.

Para López et al. (2008), la ecuación (37a) se explica a través de tres términos<sup>47</sup>: (a)  $\beta_j^{pl} (\bar{X}_j^{pr} - \bar{X}_j^{pl})$  corresponde a la contribución de las diferencias en las variables que determinan el nivel salarial (capital humano, características de la empresa, etc.), mientras que (b)  $\bar{X}_j^{pr} (\beta_j^{pr} - \beta_j^{pl})$  hace referencia a la contribución de las diferencias en los rendimientos estimados para esas variables y (c) el tercer término,  $(\bar{\lambda}^{pr} \sigma^{pr} - \bar{\lambda}^{pl} \sigma^{pl})$ , es el debido a la existencia del proceso de selección.

Si se asume que la estructura salarial no discriminatoria es la del sector privado, entonces la expresión (37a) se expresa como sigue:

$$\sum_j \beta_j^{pr} \bar{X}_j^{pr} - \sum_j \beta_j^{pl} \bar{X}_j^{pl} = \sum_j \beta_j^{pr} (\bar{X}_j^{pr} - \bar{X}_j^{pl}) + \sum_j \bar{X}_j^{pl} (\beta_j^{pr} - \beta_j^{pl}) + (\bar{\lambda}^{pr} \sigma^{pr} - \bar{\lambda}^{pl} \sigma^{pl}) \quad (38a)$$

Similarmente a la ecuación (37a) se interpreta y se descompone la expresión (38a).

Finalmente, dado que el presente trabajo de investigación aplica modelos econométricos basados en supuestos para explicar la realidad, existen ciertas

<sup>46</sup> La ecuación (33) es el resultado de la diferencia entre las ecuaciones (32c) y (32d).

<sup>47</sup> Los primeros dos términos corresponden a los de la descomposición de Oaxaca-Blinder (1973) clásica en ausencia del proceso de selección ya mencionados en la interpretación de la ecuación (33).

limitaciones metodológicas, entre ellas se encuentra el sesgo de selección muestral que es corregido por el método planteado por Heckman (1979); por lo tanto, no se realizaron aproximaciones a través del estimador de Variables Instrumentales u otras opciones de estimación debido fundamentalmente a la disponibilidad de datos. En este sentido, no es posible corregir el sesgo de endogeneidad de la educación ya que es difícil encontrar variables que no estén correlacionadas con el término de error y a su vez estén correlacionadas con la variable educación. Otro aspecto que no se considera en este trabajo es la denominada pre-discriminación salarial relacionada principalmente con el “background” familiar<sup>48</sup> debido igualmente a que en el Ecuador no existe esta información.<sup>49</sup>

Asimismo, algunas investigaciones no efectúan la estimación por MCO de dos etapas desarrollado por Heckman (1979) ya que no encontraron una variable que sea correlacionada con el sector de empleo y que no afecte a los ingresos del trabajador; en consecuencia, para resolver el sesgo de selección realizan simulaciones asumiendo diferentes niveles de correlación entre el salario y las regresiones de selección. Estas simulaciones no se realizan en el presente trabajo de investigación.

Adicionalmente, entre otras de las limitaciones del presente documento es que no se realiza la descomposición que considera tres componentes (“threefold decomposition”) de las diferencias salariales entre los sectores económicos, pues solo se realiza la descomposición en dos componentes denominada “twofold decomposition”. Además, estas formas de descomposición salarial tienen la limitación de que se debe tomar como estructura salarial no discriminatoria a la de un grupo de referencia, lo que conlleva a ciertos problemas de medición que se mencionarán más adelante; por tanto, Blinder-Oaxaca (1973) han propuesto un modelo más general para solucionar dicho problema; no obstante, el presente trabajo de investigación no estima tal modelo sino que se propone para futuras investigaciones.

---

<sup>48</sup> Considera si el padre o la madre se ha educado, los ingresos de los padres, etc.

<sup>49</sup> Para una explicación más extensiva véase Cain (1986) y para el caso ecuatoriano Torresano (2009).

Por otra parte, la mayoría de trabajos empíricos que tienen como objetivo el análisis empírico de las diferencias salariales entre el sector público y privado obtienen este diferencial según el género ya que la participación en el mercado de trabajo es muy diferente entre hombres y mujeres. En este sentido, la presente investigación tampoco realizará dichas descomposiciones debido a que estos aportes pueden ser planteados en futuros trabajos.



## **CAPITULO IV**

### **RESULTADOS**

#### **4.1. DESCRIPCIÓN DE LOS DATOS**

Para el análisis de los diferenciales salariales entre los sectores público y privado, esta investigación utiliza los datos del Instituto Nacional de Estadística y Censos (INEC) de la Encuesta Nacional de Empleo, Desempleo y Subempleo (ENEMDU) del mes de diciembre de 2011; ya que según el INEC, dicha encuesta tiene como objetivo, “proporcionar información sobre el Mercado Laboral Ecuatoriano, a través de la recolección de datos, con periodicidad mensual en las principales ciudades y trimestral en el total nacional urbano, siguiendo las normas y recomendaciones internacionales de la OIT”. Por tanto, la ENEMDU facilita datos importantes para el estudio tales como salarios, sector público o privado al que pertenecen los trabajadores, actividad laboral, género, nivel de educación, etc. En conclusión, dicha encuesta proporciona características socio demográficas que pueden explicar las diferencias salariales entre los trabajadores.

La ENEMDU tiene cobertura nacional y el levantamiento de información es realizado en el área urbana y rural. Según el INEC, la muestra de esta investigación está formada por todos los hogares particulares dentro del territorio ecuatoriano, incluido las zonas “no delimitadas”; sin embargo se excluyen: la Región Insular, la población en viviendas colectivas, viviendas flotantes y sectores con población indigente.

El INEC indica que para la ENEMDU se aplica una submuestra de la Muestra Maestra; esta muestra se construye en base de la cartografía del VI Censo de

Población y Vivienda del 2001 (CPV-2001) con la asesoría del Buró del Censo de los Estados Unidos de América (BUCEN) para ser aplicado al Sistema Integrado de Encuestas de Hogares<sup>50</sup> (SIEH).

En el diseño de la Muestra Maestra, el INEC define *UPMs* (Unidades Primarias de Muestreo) como un conjunto de sectores censales. Las *UPMs*, se seleccionan con probabilidad proporcional al tamaño en la primera etapa y dentro de cada *UPM* se selecciona un sector con probabilidad proporcional al tamaño (PPT) en la segunda etapa. El INEC considera que el tamaño promedio para una *UPM* urbana es de 450 viviendas y para el área rural de 320 viviendas.

Según el INEC, la unidad primaria de selección es la (*UPM*); la unidad secundaria son los sectores seleccionados (uno dentro de cada *UPM*); la tercera unidad es la vivienda. Como se mencionó en el párrafo anterior, las *UPMs* se seleccionan con probabilidad proporcional al tamaño (PPT). Por esta razón la ENEMDU se convierte en una muestra probabilística y trietápica:

#### ❖ **Primera Etapa**

Unidades Primarias (*UPM*): *UPM*'s.

Número de *UPM*'s en la muestra maestra: 2.094 *UPM*'s.

#### ❖ **Segunda Etapa (USM)**

Unidades: 7.409 sectores censales en la muestra maestra.

#### ❖ **Tercera Etapa**

Unidades: las viviendas ocupadas.

Fijación uniforme: 12 viviendas ocupadas por sector censal.

---

<sup>50</sup> Según el INEC, el *Sistema Integrado de Encuestas de Hogares – SIEH*, constituye un instrumento de planeación básica estratégica, que permite bajo un método integrado, desarrollar la producción de información estadística sociodemográfica y económica, en el marco de un conjunto de encuestas de hogares que se vinculan entre sí, a través de un enfoque conceptual y metodológico común.

La muestra para la ENEMDU trimestral según el INEC es de 573 sectores muestrales urbanos y comprende 69.653 encuestados de forma aleatoria. Pero para este estudio se tomó en consideración una submuestra que contiene 5.542 individuos con las siguientes características: hombres o mujeres con distinta identificación étnica, cuya edad oscila entre 10 y 65 años, que forman parte de la población ocupada, que muestran un registro de ingresos<sup>51</sup>, horas trabajadas<sup>52</sup> y otras características laborales y demográficas; además, en esta sub muestra se excluye a los individuos que no forman parte del sector formal y pertenecen a una organización cuya rama de actividad es la Agricultura, Extracción de Petróleo, Manufactura, y Generación Eléctrica<sup>53</sup>; también, se excluyen a los trabajadores por cuenta propia, trabajadores del hogar no remunerado, trabajadores no del hogar no remunerado, ayudante no remunerado de asalariado/jornalero; se descartan a estos individuos ya que no son trabajadores dependientes, no reciben un salario fijo o simplemente no perciben ninguna remuneración.

Entre las variables que se construyen en base a la ENEMDU para la aplicación de los modelos de ingresos planteados para la población ocupada son:

- i. **Logaritmo del ingreso corriente por hora proveniente del trabajo:** el cual es considerado como variable dependiente y es calculado por el logaritmo neperiano de la división de los ingresos mensuales laborales corrientes, tanto de actividades principales y secundarias y las horas mensuales trabajadas por

---

<sup>51</sup> Se consideran a los ingresos mayores que cero y menores-iguales a 10000.

<sup>52</sup> El INEC se basa en definiciones internacionales por lo general planteadas por la Organización Internacional del Trabajo (OIT), por tanto considerando la definición de Ocupado que toma en cuenta a las personas que en la semana de referencia han trabajado al menos una hora, la sub muestra en la que se basa el presente trabajo de investigación incluye solamente a los individuos que trabajan de 1 a 52 horas semanales. El justificativo de hasta 52 horas semanales hace referencia al Código de Trabajo Ecuatoriano, capítulo V, artículo 47 en el que se manifiesta que: “La jornada máxima de trabajo será de ocho horas diarias, de manera que no exceda de cuarenta horas semanales, salvo disposición de la ley en contrario.” y en el artículo 55 literal 1 que indica: “Las horas suplementarias no podrán exceder de cuatro en un día, ni de doce en la semana”.

<sup>53</sup> Según Carrillo (2004) se deben incluir a las personas que forman parte del sector formal y se deben excluir a los individuos que pertenecen a una organización cuya actividad esté vinculada con la Agricultura, Extracción de Petróleo, Manufactura, y Generación Eléctrica ya que la encuesta puede clasificar a aquellos empleados de empresas de constitución privada, pero de propiedad del Gobierno, como trabajadores del sector privado.

empleados dependientes. Las horas mensuales trabajadas se calculan de la misma manera que Torresano (2009), por tanto se multiplican las horas semanales efectivas de trabajo por 4.3 semanas del mes.

En cuanto a la definición de ingreso corriente planteada por el INEC se tiene que: son las percepciones monetarias y/o en especie que recibe la persona, en forma habitual y son percibidas a intervalos regulares de tiempo (diario, semanal, quincenal, mensual). Además, según el INEC, los ingresos laborales corrientes son distintos para las personas asalariadas o dependientes y para los trabajadores independientes, estos últimos no son tomados en cuenta en la sub muestra seleccionada; por tanto, i) los ingresos totales de las personas ocupadas que sean asalariadas incluye: el sueldo salarial básico unificado, el ingreso que recibió por concepto de componente salarial en el proceso de incorporación a las remuneraciones, subsidios, gastos de representación, bonificaciones, utilidades, vacaciones, décimo tercero o cuartos sueldos y otros beneficios. Además incluye el ingreso que recibió en el último mes por horas extras, comisiones o propinas y los descuentos que tuvo por pago al seguro social, impuestos a la renta, etc. Finalmente se incluye el ingreso que recibió por alimentos y vivienda; y ii) los ingresos totales de las personas ocupadas que trabajan por cuenta propia incluyen ingresos por venta de bienes o prestación de servicios, el consumo hecho por miembros de familia de bienes o servicios del negocio, y de esto se resta lo que se gasta en total en el último mes para el funcionamiento del negocio.

- ii. **Variables de capital humano:** Estas variables hacen referencia a los conocimientos y habilidades que poseen las personas, que las capacita con el objetivo de que sean competentes para realizar labores productivas. En la presente investigación estas variables se consideran exógenas y son:

❖ **Educación:** Años de educación aprobados por el individuo declarados en la encuesta.

- ❖ **Experiencia potencial:** Según Altonji y Blank (1999), la experiencia potencial para cada individuo se expresa en años y puede ser calculada como el resultado de la diferencia de la edad - los años de educación – 6 y bajo el supuesto de que las personas están trabajando en su edad adulta cuando no están en la escuela.
  - ❖ **Nivel de educación:** Corresponde a cuatro variables ficticias que permitirán obtener los rendimientos para cada nivel y estas son: **sin estudios, educación primaria, secundaria, y superior**. Las personas que declararon tener un nivel de instrucción primaria, centro de Alfabetización o Jardín de Infantes forman parte del nivel correspondiente a educación primaria; quienes declararon un nivel de educación básica, secundaria o educación media/bachillerato se ubican en el nivel de educación secundaria y para el nivel de instrucción superior corresponde a las categorías de educación superior no universitaria, superior universitaria y post grado.
- iii. **Variables de características personales:** Son variables que indican cualidades que permiten identificar el carácter propio y específico de las personas. A continuación se detallan estas variables consideradas independientes por los modelos econométricos planteados.
- ❖ **Mujer:** Es una variable dicotómica de género que toma el valor de 1 si el individuo es mujer y 0 si es hombre.
  - ❖ **Indígena:** Hace referencia a una variable indicatriz que es igual a 1 cuando el individuo se auto identifica como indígena y 0 caso contrario.

La variable Indígena se crea en base a la variable etnia proporcionada por la ENEMDU; por tanto, es importante mostrar la definición de etnia, ya que a lo largo del tiempo han existido largos debates sobre las definiciones en las encuestas de hogares de raza y etnia.

Según Mejía et al. (2000) **etnia** se la define como la pertenencia o identificación con un grupo que se diferencia del resto de la población no sólo por sus rasgos físicos sino más bien por su idioma, tradiciones, etc., que los identifica como "población indígena". Mientras que **raza** se entiende como la pertenencia o identificación con un grupo que se distingue del resto de la población sólo en base a sus rasgos físicos.

En base a la disponibilidad de datos, las variables lenguaje hablado, autopercepción y la concentración geográfica permiten identificar a los miembros del grupo indígena (considerado en este trabajo como minoritario). Sin embargo, Torresano (2009) indica que en la Convención de Indígenas y Pueblos Tribales de la "Organización Internacional del Trabajo" (OIT) en el año 1989 (No. 169) se estableció que la auto identificación como indígena o tribal debería ser considerada como un criterio fundamental para la determinación de los grupos por etnia. En tal sentido, en Ecuador la encuesta ENEMDU provee de información respecto de la auto identificación étnica; por consiguiente, en el presente trabajo se considera a una persona como parte de la población indígena, siempre y cuando se haya auto identificado con dicho grupo.

- ❖ **Relación de parentesco en la familia:** Corresponde a una variable binaria que toma el valor de 1 en caso de que el individuo sea jefe de familia y 0 en caso de que la relación de parentesco del individuo en la familia sea conyugue, hijo, yerno o nuera, nieto, padres, suegros, otros parientes, empleada doméstica u otros no parientes. Cabe mencionar que esta variable es seleccionada, pues según el INEC, los miembros del hogar pueden reconocer como jefe a un integrante ya sea por una mayor responsabilidad en las decisiones familiares, por prestigio, relación familiar o de parentesco, por tradiciones culturales o por razones económicas; por tanto, esta última característica es una aproximación del grado de responsabilidades que tiene

el individuo; entonces, se esperaría que la probabilidad de que estos individuos tengan un mayor ingreso sea alta. En consecuencia, esta variable tendría influencia en el análisis de las diferencias salariales.

- ❖ **Situación migratoria:** Se introduce una variable indicatriz que toma el valor de 1 si el individuo ha declarado que migró de su ciudad o región original, 0 en caso contrario.
- ❖ **Estado civil:** Se genera una variable dicotómica igual a 1 en el caso en el que el individuo haya declarado ser casado o en unión libre y 0 si declaró ser soltero, divorciado, separado o viudo.
- ❖ **Localización geográfica:** Se introducen siete variables indicatrices que representan las zonas norte, sur y centro de la regiones Sierra, Costa y Amazonía, además se crea una variable binaria que es igual a 1 en el caso de que los individuos estén ubicados en la ciudad de Guayaquil. Es importante notar que no se considera a la región insular como ya se mencionó anteriormente.

Las variables que se están tomando en consideración se basan en la información que la ENEMDU proporciona; sin embargo, se podría introducir variables adicionales en base a investigaciones internacionales dependiendo de la disponibilidad de datos y resultados de los modelos; así por ejemplo Cain (1986) propone las siguientes variables: educación, edad, raza, habilidad mental, capacitación formal, experiencia actual en el mercado laboral, variables que representen la experiencia en el mercado laboral como la experiencia potencial<sup>54</sup>, estado civil, salud, horas de trabajo (anual, semanal, jornada completa, tiempo parcial), permanencia (tiempo de servicio en el empleo actual), tamaño de la ciudad de residencia, región de residencia, background familiar (educación/ocupación/ingreso del padre, número de hermanos, historia

---

<sup>54</sup> En la presente investigación, la experiencia potencial se obtiene de la resta de la edad del individuo menos los años de educación y menos 6.

migratoria, etnicidad, etc.), registro de ausentismo, calidad de la educación, doble carga (número de hijos, límites en hora de ubicación, planes de dejar el trabajo por razones distintas de la formación, etc.), urbano/rural, volumen del negocio, ocupación, prestigio ocupacional, tipo de trabajador (por cuenta propia, público o privado de sueldos y salarios), industria, sindicatos, tipo de empleador (privado/público, genero segregado/integrado, tamaño de la fuerza de trabajo), estado de supervisión, porcentaje de mujeres en el grupo de trabajo, mediana del ingresos de los hombres titulares, condiciones del mercado laboral local, duración del viaje al trabajo, tiempo de viaje de trabajo, estatus de veterano, estatus migratorio.

Para Torresano (2009), tanto el Modelo Aditivo, Multiplicativo y los modelos de Descomposición de ingresos entre sectores económicos pueden presentar en sus estimaciones por MCO un problema de sesgo de selección muestral; por tanto, se plantea la corrección de dicho sesgo utilizando la metodología de Heckman (1979) que sugiere la estimación de un modelo probit de participación laboral mediante el cual se obtiene el ratio de Mills, seguidamente este ratio se introduce como un nuevo regresor en los modelos de ingresos mencionados con el objetivo de resolver el problema. Para la aplicación del modelo probit, Rama (2003) sugiere la utilización de regresores determinantes de la probabilidad de participar o no en el mercado laboral; estos regresores constituyen un conjunto de características individuales tales como: edad, género, localización geográfica (páramo), años de escolaridad (en grupos: 1-5, 6-8 y 9-10 años), Certificado General de Educación<sup>55</sup> en Nivel Ordinario y en Nivel Avanzado (GCE, O/L y A/L)<sup>56</sup>, título universitario o de postgrado, formación profesional (en años), esposa o esposo como jefe de hogar, hijo o hija como jefe de hogar, otros no como jefe de hogar y tipo de distrito (urbanos, rural o estados agrícolas). Además, en el modelo probit propuesto por Rama (2003), la variable dependiente toma el valor de 1 si el individuo es desempleado y cero si esta empleado, esto equivale a decir en el presente trabajo de investigación que la

---

<sup>55</sup> El Certificado General de Educación (GCE) es un título académico que los consejos de examinación en el Reino Unido y en algunas de las antiguas colonias británicas o países mancomunados, en particular Sri Lanka y Singapur, confieren a los estudiantes.

<sup>56</sup> Por sus siglas en inglés.



variable endógena es una variable indicatriz que hace referencia a la decisión de participar o no en el mercado laboral (toma el valor de 1 si los trabajadores pertenecen al conjunto de la población ocupada y 0 en caso contrario). Por otra parte, Torresano (2009) indica que la literatura económica sugiere la utilización de variables exógenas tales como: género, experiencia y educación; sin embargo, factores ambientales, sociales y personales condicionan la participación laboral; por consiguiente, se generan variables de características personales y del hogar a través de la información que proporciona la ENEMDU (entre estas variables se encuentran el estado civil, relación de parentesco, número de miembros del hogar, situación migratoria y localización geográfica).

#### **4.2. ANÁLISIS DESCRIPTIVO DE LOS DIFERENCIALES SALARIALES POR SECTOR LABORAL**

En el Ecuador, según el VII Censo de Población y el VI de Vivienda realizado en 2010, la población total oficial en el país es 14'483.499 y hay 4'654.054 viviendas, pero para el 2011 la población incrementa a 15.2 millones de acuerdo a las proyecciones poblacionales realizadas por el INEC.

Para el análisis descriptivo de los diferenciales salariales por sector económico, este estudio consideró una sub muestra de 5.542 individuos con las características ya mencionadas en el apartado anterior. Por tanto, la Tabla- 1 resume las características principales de la sub muestra respecto a cada uno de los grupos poblacionales (sector público y privado). Así, el 44,41% de los individuos que pertenecen a la sub muestra forman parte del sector público y el 55,59% se encuentran ocupados en el sector privado. Con respecto al género, en el sector privado existen un mayor número de hombres empleados, pues representan el

58,65%; mientras que las mujeres representan el 41,35%; pero en el sector público la situación es diferente ya que el 50,91% son mujeres y el 49,09% son hombres.

En cuanto al área de residencia de los individuos se puede observar que tanto el sector público como el sector privado están conformados por trabajadores que viven en su mayoría en las zonas urbanas del país (en el sector privado el 84,52% de los individuos y en el sector público el 83,67% de las personas). Desde la perspectiva étnica, la población indígena corresponde a la minoría en ambos sectores (público y privado), pues el 1,95% del sector privado es indígena y el 98,05% pertenece a la población auto identificada como no indígena; igualmente, en el sector público el 3,82% se auto identifica como indígena y el 96,18% se considera no indígena.

La Tabla- 1 indica que los individuos que forman parte del sector privado son relativamente jóvenes, ya que aproximadamente el 45% de la población se encuentra entre los 10 y 30 años de edad; esta cifra incrementa considerablemente para los individuos entre los 10 y 50 años, pues su valor aproximado es de 88%; de igual manera, en el sector público se observa que aproximadamente el 20% de la población tiene entre 10 y 30 años; así también, en este mismo sector el 70% de los individuos se encuentran entre los 10 y 50 años de edad. Es importante notar en base a estos resultados que el sector privado tiene un mayor número de empleados jóvenes con respecto al sector público; además, la evidencia también es que el 12% de los individuos que forman parte del sector privado tienen entre 50 y 65 años de edad, mientras que en el sector público en este mismo rango se encuentra aproximadamente el 30% de las personas.

Respecto al nivel de instrucción en ambos sectores, los resultados muestran que el sector público en su mayoría, ha abarcado a empleados con educación superior ya que estos representan el 70% de los individuos; de igual forma, el sector privado en su mayoría tiene a su cargo personas con instrucción superior pero en un menor porcentaje (44,89%). Seguidamente, el sector privado y público abarcan a trabajadores que tienen instrucción secundaria, con el 38,43% y el 23,12% de los

individuos respectivamente. Es fácil observar que los dos sectores reclutan a personas con mayores niveles de educación que a individuos con menores niveles.

Con el objetivo de analizar la condición de actividad de cada sector, a continuación se muestra en la siguiente figura la clasificación de la población por actividad conforme el INEC:

**Gráfico- 1 Clasificación de la Población por Actividad**



Fuente: Instituto Nacional de Estadísticas y Censos del Ecuador (INEC)  
Elaboración: La autora

La condición de actividad<sup>57</sup> de los grupos muestra que en el sector público la mayoría de individuos se encuentran ocupados plenamente (65,93%), seguido por las personas que pertenecen al grupo de otras formas de sub-empleo (30,81%). Por otra parte, haciendo referencia al sector privado, la Tabla- 1 muestra que la mayor parte de los empleados se encuentran en la inactividad (48%); no obstante, aunque en un

<sup>57</sup> El análisis descriptivo de la condición de actividad se realiza en base de toda la población de la ENEMDU mas no de la sub muestra escogida.

menor porcentaje también existen individuos que están en otras formas de subempleo (27,38%) y son ocupados plenos (17,6%).

En relación a las características de los puestos de trabajo (grupo de ocupación), un alto porcentaje de individuos del sector público son “Profesionales científicos e intelectuales” (45,35%), de la misma manera, en este sector económico, el 15% son “Técnicos y profesionales de nivel medio” y un 10 % son “Empleados de oficina”, al igual que el 10% que son “Trabajadores de los servicios y comerciantes”. Por otra parte, el sector privado se caracteriza por tener un alto porcentaje de trabajadores que forman parte del grupo de “Trabajadores de los servicios y comerciantes”, siendo el 21,36% de personas que conforman dicho grupo; en segundo lugar se encuentran los “Empleados de oficina” con el 18,37% de la sub muestra, seguidamente los “Técnicos y profesionales de nivel medio” son el 14,51% de individuos del sector privado, después el 13,63% de los trabajadores son “Profesionales científicos e intelectuales” y el 12,69% son “Trabajadores no calificados”.

Finalmente, las estadísticas indican que existe una diferencia promedio de 2 horas semanales efectivas trabajadas entre ambos sectores económicos, siendo el sector público el que presenta más horas semanales trabajadas en promedio (42,07 horas) con respecto al sector privado (40,11 horas); adicionalmente se puede notar que el promedio de las horas semanales trabajadas tanto del sector público como del sector privado concuerdan con lo estipulado en el Código de Trabajo Ecuatoriano que especifica que no se debe exceder las 40 horas de trabajo semanales salvo la ley especifique lo contrario. Por otra parte, la diferencia de la media de los ingresos laborales corrientes de actividad principal y secundaria por hora es de 2,77 dólares a favor del sector público.

**Tabla- 1 Características de una Sub muestra de la ENEMDU por Sector Económico (Porcentajes)**

VARIABLES	SECTOR PRIVADO	SECTOR PÚBLICO
<b>Total</b>	55,59	44,41
<b>Género</b>		
Mujer	41,35	50,91
Hombre	58,65	49,09
<b>Área de Residencia</b>		
Urbana	84,52	83,67
Rural	15,48	16,33
<b>Etnia</b>		
Indígenas	1,95	3,82
No Indígenas	98,05	96,18
<b>Grupos de Edad</b>		
Menores a 20	6,91	1,06
21-30	38,43	19,38
31-40	26,55	23,97
41-50	16,13	25,44
51-60	10	24,95
61-65	1,98	5,2
<b>Nivel de Instrucción</b>		
Ninguno	0,55	0,24
Educación Primaria	16,13	6,54
Educación Secundaria	38,43	23,12
Educación Superior	44,89	70,09
<b>Condición de Actividad</b>		
Ocupados plenos	17,6	65,93
Ocupados no clasificados	0,2	0,99
Subempleo visible	4,71	0,61
Otras formas de subempleo	27,38	30,81
Desempleo abierto	1,33	1,32
Desempleo oculto	0,78	0,34
Inactivos	47,99	0
<b>Grupo de Ocupación</b>		
Fuerzas Armadas	0	2,56
Personal direct./admin. pública y empresas	2,63	3,9
Profesionales científicos e intelectuales	13,63	45,35
Técnicos y profesionales de nivel medio	14,51	14,95
Empleados de oficina	18,37	10,16
Trabajad. de los servicios y comerciantes	21,36	10,08
Trabajad. calificados agropecuarios y pesqueros	0,68	0,45
Oficiales operarios y artesanos	9,48	1,91
Operadores de instalac. máquinas y montad.	6,65	4,47
Trabajadores no calificados	12,69	6,18
<b>Otras características laborales</b>		
Horas semanales efectivas trabajadas	40,11	42,07
Ingresos Laborales Corrientes de Actividad principal y secundaria por hora	1,8	4,57

Fuente: Instituto Nacional de Estadísticas y Censo del Ecuador (INEC) - Encuesta de Empleo y Desempleo - Diciembre 2011

Elaboración: La autora

Notas:

(1) Para conocer las definiciones de las Condiciones de Actividad establecidas por el INEC, véase Aspectos Metodológicos de ENEMDU. Estas estadísticas no toma en cuenta solo a la sub muestra seleccionada, sino que considera a toda la población ecuatoriana.

(2) Los datos están expresados en dólares corrientes.

### 4.3. ESTIMACIONES: MODELOS ADITIVOS, MULTIPLICATIVOS Y DESCOMPOSICIÓN SALARIAL POR SECTOR

#### 4.3.1 ANÁLISIS DE DISCRIMINACIÓN SALARIAL: MODELOS ADITIVOS Y MULTIPLICATIVOS

En el presente apartado se analiza los resultados de las estimaciones de las ecuaciones de ingresos controladas por sector laboral para la sub muestra seleccionada (5.542 individuos); estas estimaciones se realizan por Mínimos Cuadrados Ordinarios (MCO) y por el método de sesgo de selección muestral propuesto por Heckman (1979); siendo este último el que incluye en la sub muestra a la población desocupada ya que como se mencionó anteriormente, la decisión de participar o no en el mercado laboral puede ser endógena.

En este sentido, en la Tabla- 2, los resultados muestran que el coeficiente del regresor denominado cociente o ratio de Mills es significativo en todos los modelos corregidos por sesgo de selección muestral; es decir, se rechaza la hipótesis nula de ausencia de sesgo de selección ( $\sigma = 0$ ); por consiguiente, los modelos corregidos son más apropiados y serán analizados en los párrafos posteriores.

El modelo aditivo corregido por el método de Heckman (1979) es así:

$$\ln(\text{ingreso}_{\text{hora}}) = \beta_0 + \beta_1 \text{educación} + \beta_2 \text{experiencia} + \beta_3 \text{experiencia}^2 + \beta_4 \text{público} + \beta_5 \text{mujer} + \beta_6 \text{indígena} + \beta_7 \text{jefe}_{\text{hogar}} + \beta_8 \text{migración} + \beta_9 \text{sierra}_{\text{norte}} + \beta_{10} \text{sierra}_{\text{centro}} + \beta_{11} \text{sierra}_{\text{sur}} + \beta_{12} \text{costa}_{\text{norte}} + \beta_{13} \text{amazonía} + \beta_{14} \text{guayaquil} + \sigma \hat{\lambda}$$

Donde,  $\hat{\lambda}$  es la inversa del coeficiente de Mills y  $\sigma$  es la covarianza entre el término de error de la ecuación de ingresos y participación laboral.

Pero es el modelo multiplicativo el cual se va a analizar puesto que se deriva del modelo aditivo e incluye una variable adicional que es la interacción entre las variables público y experiencia; por tanto, el modelo multiplicativo corregido por el sesgo de selección muestral que se construyó para el análisis de discriminación salarial es el siguiente:

$$\begin{aligned} \ln(\text{ingreso}_{\text{hora}}) &= \beta_0 + \beta_1 \text{educación} + \beta_2 \text{experiencia} + \beta_3 \text{experiencia}^2 + \beta_4 \text{público} \\ &+ \beta_5 \text{mujer} + \beta_6 \text{indígena} + \beta_7 \text{jefe}_{\text{hogar}} + \beta_8 \text{migración} + \beta_9 \text{sierra}_{\text{norte}} \\ &+ \beta_{10} \text{sierra}_{\text{centro}} + \beta_{11} \text{sierra}_{\text{sur}} + \beta_{12} \text{costa}_{\text{norte}} + \beta_{13} \text{amazonía} \\ &+ \beta_{14} \text{guayaquil} + \beta_{14} \text{público} \times \text{experiencia} + \sigma \hat{\lambda} \end{aligned}$$

Sea  $\sigma$  y  $\hat{\lambda}$ , los términos ya definidos en el modelo aditivo.

Las estimaciones del modelo multiplicativo<sup>58</sup> que se presentan en la Tabla- 2 muestran que la hipótesis planteada para el presente trabajo de investigación es verdadera ya que existe premio salarial entre el Sector Público y Privado. En particular, los resultados evidencian discriminación salarial a favor del sector público y en contra del sector privado, pues el coeficiente de la variable indicatriz que representa al sector económico (que toma el valor de 1 si el individuo pertenece al sector público y 0 caso contrario) es positiva y significativa. Por lo tanto, se tiene que el sector público recibe un 28% adicional en sus ingresos con respecto al sector privado (manteniendo constantes el resto de variables de control)<sup>59</sup>.

Por otra parte, se identifica que el coeficiente de la variable educación que expone los rendimientos de la educación es positivo, significativo e igual a 6% (en el modelo no corregido este resultado es de aproximadamente 7%); en otras palabras, por cada año adicional de educación y si los demás regresores permanecen constantes, los ingresos podrían incrementar en promedio 6%. Otra de las variables que representan al capital humano es la experiencia, cuyo coeficiente también es positivo y

<sup>58</sup> La validación del modelo multiplicativo se puede observar en el Anexo- A.

<sup>59</sup> Este resultado es similar al de MCO.

significativo, y muestra que por cada año adicional de experiencia (manteniendo el resto de factores observables constantes), el ingreso individual puede incrementar en aproximadamente 2% para las estimaciones por MCO y por el método de Heckman (1979). A diferencia de la experiencia, la variable experiencia al cuadrado es negativa pero significativa; por tanto, se verifica la concavidad hacia abajo de la función de ingresos con respecto de la experiencia ya que los individuos reciben su paga en forma creciente hasta llegar a un máximo de la experiencia laboral y luego estos ingresos empiezan a decrecer.

En relación a los rendimientos de educación de este estudio con respecto a los resultados de Carrillo (2004) al parecer no ha existido un cambio significativo pues su porcentaje promedio se mantiene en 7%.

Con respecto a las variables relacionadas con las características individuales se observa que aparentemente existe discriminación salarial contra las mujeres e indígenas (los coeficientes de ambas variables ficticias son negativos y significativos). Particularmente, las mujeres reciben un salario menor en aproximadamente 12% en relación a la remuneración que recibe el género masculino; de igual forma, los individuos auto identificados como indígenas obtienen 12% menos en sus ingresos con respecto a la población no indígena<sup>60</sup>. También, las estimaciones indican que las personas que se consideran jefe de familia (manteniendo el resto de variables exógenas constantes) pueden recibir un pago adicional de aproximadamente 5%; de igual manera, tener una condición migratoria (manteniendo los demás regresores constantes), podría favorecer a recibir mayores ingresos en promedio 6%. Finalmente, las variables relacionadas con localización geográfica de los individuos, muestran que tomando como referencia a la región costa, ubicarse en las demás zonas geográficas es favorable para recibir un mayor salario; pues los coeficientes son positivos y significativos: Sierra Norte (7%), Sierra Centro (7%), Sierra Sur (10%) y Amazonía (19%). Por otra parte, habitar en la ciudad de Guayaquil (considerada como el puerto principal y la ciudad con el mayor número

---

<sup>60</sup> En el caso de ambos grupos se mantienen fijos el resto de regresores observables.



de habitantes del Ecuador) permitiría obtener salarios mayores en un 13%, considerando que este valor es positivo y significativo.

En cuanto a la variable de interacción del modelo multiplicativo, los resultados indican que tener simultáneamente mayor experiencia y trabajar en el sector público (manteniendo las demás variables constantes), favorecería para obtener mayores rendimientos laborales.

En otro sentido, los modelos aditivos que representan al sector público y privado por separado se describen a continuación:

$$\begin{aligned} \ln(\text{ingreso}_{\text{hora}}^{pr}) &= \beta_0^{pr} + \beta_1^{pr} \text{educación}^{pr} + \beta_2^{pr} \text{experiencia}^{pr} + \beta_3^{pr} \text{experiencia}^{2pr} \\ &+ \beta_4^{pr} \text{público}^{pr} + \beta_5^{pr} \text{mujer}^{pr} + \beta_6^{pr} \text{indígena}^{pr} + \beta_7^{pr} \text{jefe}_{\text{hogar}}^{pr} \\ &+ \beta_8^{pr} \text{migración}^{pr} + \beta_9^{pr} \text{sierra}_{\text{norte}}^{pr} + \beta_{10}^{pr} \text{sierra}_{\text{centro}}^{pr} \\ &+ \beta_{11}^{pr} \text{sierra}_{\text{sur}}^{pr} + \beta_{12}^{pr} \text{costa}_{\text{norte}}^{pr} + \beta_{13}^{pr} \text{amazonía}^{pr} \\ &+ \beta_{14}^{pr} \text{guayaquil}^{pr} + \sigma^{pr} \hat{\lambda}^{pr} \end{aligned}$$

$$\begin{aligned} \ln(\text{ingreso}_{\text{hora}}^{pl}) &= \beta_0^{pl} + \beta_1^{pl} \text{educación}^{pl} + \beta_2^{pl} \text{experiencia}^{pl} + \beta_3^{pl} \text{experiencia}^{2pl} \\ &+ \beta_4^{pl} \text{público}^{pl} + \beta_5^{pl} \text{mujer}^{pl} + \beta_6^{pl} \text{indígena}^{pl} + \beta_7^{pl} \text{jefe}_{\text{hogar}}^{pl} \\ &+ \beta_8^{pl} \text{migración}^{pl} + \beta_9^{pl} \text{sierra}_{\text{norte}}^{pl} + \beta_{10}^{pl} \text{sierra}_{\text{centro}}^{pl} \\ &+ \beta_{11}^{pl} \text{sierra}_{\text{sur}}^{pl} + \beta_{12}^{pl} \text{costa}_{\text{norte}}^{pl} + \beta_{13}^{pl} \text{amazonía}^{pl} \\ &+ \beta_{14}^{pl} \text{guayaquil}^{pl} + \sigma^{pl} \hat{\lambda}^{pl} \end{aligned}$$

Donde el superíndice  $pl$  hace referencia al Sector Público, y el superíndice  $pr$  representa al Sector Privado y los términos  $\sigma$  y  $\hat{\lambda}$  se definieron en la representación del modelo aditivo.

Adicionalmente, la Tabla- 2 exhibe los resultados de las estimaciones de las ecuaciones de ingresos por separado para cada sector laboral<sup>61</sup> (público y privado). De estas estimaciones se observa que tanto los trabajadores públicos y privados tienen los mismo rendimientos educativos (6%); en otras palabras, por un año adicional de educación (manteniendo constantes los demás factores), los ingresos incrementan en promedio 6% tanto para los trabajadores del sector público y de la misma manera para los trabajadores privados. Otra variable a analizar que caracteriza al capital humano es la experiencia cuyo coeficiente es significativo y positivo para ambos sectores, lo que quiere decir que por un año adicional de experiencia, con las demás variables explicativas constantes, la paga para los empleados del sector público crece en promedio 2%, este mismo porcentaje representa el incremento promedio del salario para los trabajadores del sector privado. La experiencia al cuadrado es negativa en los dos sectores económicos pues como se mencionó anteriormente existe concavidad negativa de la función de ingresos con respecto a la experiencia debido a que los salarios crecen a través del tiempo hasta que llegan a un máximo cuando empiezan a decaer<sup>62</sup>; además, el coeficiente es significativo solo para la ecuación de ingresos del sector privado y no es significativo en la ecuación del sector público, esto puede ser causa de los tipos de contratos de este último sector ya que los salarios de estos trabajadores en su mayoría crecen con el tiempo más no disminuyen.

En referencia a las características individuales, la variable género es negativa y significativa en las estimaciones de ambos sectores laborales, lo que significa que existe discriminación de género tanto en el sector público como en el privado; pero la penalidad es mayor en el sector público (los coeficientes son 11 % en el sector privado y 15% en el público). Lo mismo ocurre con la variable indígena, parece que existe mayor discriminación hacia la población indígena en el sector público que en el sector privado, pues sus coeficientes son 11% para el sector privado y 12% para el

---

<sup>61</sup> La validación de los modelos aditivos para cada sector laboral por separado se encuentra en el Anexo- B.

<sup>62</sup> Los ingresos decrecen después de cierto punto de la experiencia porque podría considerarse de que los individuos en edad adulta son menos productivos.

sector público. En relación al parentesco, un asalariado del sector privado que se considera jefe de hogar, manteniendo los demás factores constantes, puede recibir un pago adicional en promedio 9% (este valor se muestra positivo y significativo solamente en los resultados de las estimaciones de la ecuación de ingresos que corresponde al sector privado ya que en las estimaciones del sector público es positivo pero no significativo). Algo similar sucede con la situación migratoria de los trabajadores, pues ser trabajador del sector público y tener una condición migratoria, manteniendo constantes el resto de factores, favorece a recibir un salario adicional en aproximadamente 11% (este coeficiente es significativo a diferencia del coeficiente en la ecuación de ingresos del sector privado que se presenta como no significativo).

En este mismo sentido, respecto a las variables de localización geográfica se observa que considerando como categoría de referencia a la región costa, los asalariados del sector público que habitan en las zonas norte sierra, centro sierra y sur sierra (con el resto de variables exógenas constantes), reciben ingresos mayores en aproximadamente 9%, 9% y 12%, respectivamente. Similarmente, los asalariados del sector privado que se encuentran ubicados en las zona norte sierra, centro sierra y sur sierra (con el resto de variables exógenas constantes), reciben pagos mayores en aproximadamente 7%, 6% y 18%, respectivamente. Es decir, la ubicación de los individuos en estas zonas afecta positivamente en mayor proporción a los salarios de los trabajadores del sector público que a los salarios de los trabajadores privados. Continuando con el análisis de los coeficientes de las variables de localización, se evidencia que los individuos que pertenecen al sector privado y están localizados en la región amazónica (manteniendo el resto de regresores fijos), tienen un salario mayor de aproximadamente 23% con respecto a la paga que reciben los asalariados privados que se encuentran ubicados en la zona costa. En cuanto a los individuos del sector público que están localizados en la región amazónica (tomando como referencia la región costa) tienen un efecto positivo sobre los salarios (18%) pero en menor proporción en relación al efecto sobre los salarios de los empleados privados ubicados en la misma región (23%). Finalmente, tanto en el sector público como en

el privado existe un efecto positivo sobre los salarios cuando los individuos se ubican en la ciudad de Guayaquil, aunque el impacto sobre las ganancias del sector privado es mayor que el impacto sobre las ganancias del sector público (sus coeficientes son positivos y significativos, y corresponden al 14% y 11% respectivamente).

En conclusión, se evidencia discriminación salarial en contra de los trabajadores del sector privado; pues los asalariados del sector público con las mismas características de los trabajadores privados reciben un mayor ingreso. Además, los resultados evidencian que tener un año adicional de experiencia o estudios genera un efecto positivo en los salarios. De igual forma, ser jefe de hogar o tener una condición migratoria permite que los trabajadores obtengan mayores ganancias. Con respecto a las variables dicotómicas de localización geográfica, las estimaciones indican que estar ubicado en las zonas sierra norte, sierra sur, sierra centro o Amazonía con respecto a la región costa, favorece a recibir salarios más altos. Además, se evidencia discriminación en contra de las mujeres y los individuos auto identificados indígenas tanto en el modelo aditivo como en los modelos que separan a los dos sectores laborales (público y privado); sin embargo, las estimaciones indican que la penalización es mayor en el sector público en contra de estos dos grupos (mujeres e indígenas).

**Tabla- 2 Estimación de ecuaciones de ingresos poblacionales y por sector económico  
(Modelos aditivo, multiplicativo, público y privado)**

	ECUACIONES DE INGRESOS							
	MCO				Heckman			
	Aditivo	Multiplicativo	Privado	Público	Aditivo	Multiplicativo	Privado	Público
	coef/(t)	coef/(t)	coef/(t)	coef/(t)	coef/(t)	coef/(t)	coef/(t)	coef/(t)
Educación	0.070*** (36.01)	0.069*** (35.42)	0.068*** (26.72)	0.071*** (22.88)	0.063*** (19.72)	0.061*** (18.96)	0.063*** (14.91)	0.063*** (12.55)
Experiencia	0.019*** (10.54)	0.018*** (9.92)	0.018*** (7.92)	0.019*** (6.62)	0.017*** (9.11)	0.016*** (8.32)	0.017*** (7.36)	0.017*** (5.25)
Experiencia^2	-0.0001*** (-2.68)	-0.0001*** (-2.90)	-0.0001*** (-2.62)	-0.0001 (-1.15)	-0.000** (-2.49)	-0.000*** (-2.73)	-0.000** (-2.53)	-0.000 (-0.83)
Público	0.329*** (20.99)	0.284*** (11.23)			0.330*** (21.02)	0.277*** (10.91)		
Mujer	-0.127*** (-8.00)	-0.126*** (-7.96)	-0.113*** (-5.36)	-0.145*** (-5.98)	-0.122*** (-7.64)	-0.120*** (-7.55)	-0.110*** (-5.16)	-0.146*** (-6.03)
Indígena	-0.112*** (-2.62)	-0.114*** (-2.66)	-0.116* (-1.69)	-0.101* (-1.83)	-0.121*** (-2.82)	-0.124*** (-2.89)	-0.114* (-1.67)	-0.123** (-2.20)
Jefe de Hogar	0.082*** (4.67)	0.084*** (4.74)	0.114*** (4.75)	0.044* (1.68)	0.054*** (2.64)	0.052** (2.54)	0.085*** (2.72)	0.021 (0.74)
Migración	0.057*** (2.88)	0.058*** (2.94)	0.028 (1.07)	0.103*** (3.36)	0.060*** (3.05)	0.062*** (3.14)	0.036 (1.34)	0.107*** (3.50)
Sierra Norte	0.084*** (4.32)	0.083*** (4.29)	0.075*** (2.96)	0.094*** (3.08)	0.073*** (3.71)	0.071*** (3.61)	0.068*** (2.63)	0.085*** (2.75)
Sierra Centro	0.075*** (3.63)	0.074*** (3.60)	0.061** (2.00)	0.087*** (3.09)	0.071*** (3.46)	0.070*** (3.42)	0.058* (1.88)	0.086*** (3.07)
Sierra Sur	0.100*** (4.58)	0.101*** (4.60)	0.081*** (2.64)	0.123*** (3.88)	0.099*** (4.49)	0.099*** (4.51)	0.080*** (2.61)	0.121*** (3.84)
Amazonia	0.190*** (5.52)	0.189*** (5.51)	0.235*** (4.22)	0.166*** (3.75)	0.186*** (5.41)	0.186*** (5.40)	0.234*** (4.21)	0.161*** (3.63)
Guayaquil	0.131*** (5.10)	0.130*** (5.05)	0.137*** (4.45)	0.102** (2.04)	0.129*** (5.02)	0.128*** (4.96)	0.135*** (4.38)	0.107** (2.13)
Público x Experiencia		0.003** (2.30)				0.003*** (2.66)		
Constante	-0.421*** (-13.01)	-0.398*** (-11.76)	-0.386*** (-9.05)	-0.132** (-2.28)	-0.235*** (-3.10)	-0.186** (-2.38)	-0.252** (-2.45)	0.102 (0.85)
R-cuadrado	0.390	0.390	0.236	0.263	0.390	0.391	0.236	0.264
chi2					1268398	1271254	3172591	2666709
aic	8300775	8297489	4690289	3615309	8295331	8290262	4690223	3612.29
rho					-.3609354	-.4028796	-.3664103	-.3223771
sigma					.5207964	.5232101	.5271157	.5083249
lambda					-0.188***	-0.211***	-0.193*	-0.164**
estadístico t lambda					(-2.73)	(-3.03)	(-1.43)	(-2.23)

Fuente: Instituto Nacional de Estadísticas y Censo del Ecuador (INEC) - Encuesta de Empleo y Desempleo - Diciembre 2011

Elaboración: La autora

Nota: Las estimaciones son significativas cuando \* p<0.1, \*\* p<0.05, \*\*\* p<0.01

(1) En esta tabla se muestran los coeficientes y entre paréntesis se encuentra el estadístico t-student.

#### 4.3.2 DESCOMPOSICIÓN SALARIAL POR SECTOR- “TWOFOOLD DECOMPOSITION BY SECTOR”

La discriminación salarial entre grupos diferentes de trabajadores no solo puede ser analizada a través de los modelos aditivos y multiplicativos, sino también mediante la descomposición de las diferencias de ingresos de los grupos. Por tanto, se utiliza la metodología propuesta por Blinder-Oaxaca (1973) ya descrita en el capítulo anterior. Esta metodología permite analizar las diferencias salariales atribuidas a las características observables de los individuos y a la discriminación salarial que comúnmente se relaciona con las características no observables.

La Tabla- 3 expone los resultados de las estimaciones de la descomposición de las diferencias de ingresos por sector laboral (público y privado)<sup>63</sup> tanto por el método de MCO como por el método de corrección de sesgo de selección muestral planteado por Heckman (1979). Es importante reiterar que la descomposición de las diferencias salariales se puede expresar de dos formas distintas, pero en este trabajo de investigación se realiza la denominada “*Twofold decomposition*”; por consiguiente, las estimaciones de los dos componentes de las diferencias salariales asumen como estructura salarial no discriminatoria tanto al sector público ( $\beta_{pl}$ ) como al sector privado ( $\beta_{pr}$ ); puesto que como mencionan Albert et al. (1998), la estructura salarial no debe elegirse arbitrariamente sino que debe fundamentarse en la teoría.

Los resultados obtenidos indican que los asalariados del sector público reciben una prima, pues sus ingresos son mayores en promedio 67 puntos logarítmicos (o 1.87 dólares por hora) con respecto a los ingresos de los asalariados del sector privado; pero luego de corregir las diferencias salariales por el sesgo de selección muestral a través del método sugerido por Heckman (1979), esta prima disminuye a 50 puntos logarítmicos (o 1.51 dólares por hora) a favor del mismo sector laboral (público). Se puede notar también que todas las estimaciones tanto del componente debido a

---

<sup>63</sup> Estos resultados corresponden a la sub muestra seleccionada(5.542 individuos).

diferencias en las dotaciones (diferencias en características observables) como del componente atribuido a la discriminación en el mercado de trabajo (diferencias en características no observables) son negativas y significativas; no obstante, para la interpretación de este último componente se considera la subdivisión en dos partes propuesta por Jann (2008); así, la primera parte ( $U_{pr}=E(X_{pr})'\delta_{pr}$ ) mide la discriminación en contra del sector privado y la segunda parte ( $U_{pl}= -E(X_{pl})'\delta_{pl}$ )<sup>64</sup> cuantifica la discriminación a favor del sector público.

En este mismo sentido, las estimaciones muestran que si se considera como estructura salarial no discriminatoria a la del sector público, el componente atribuido a la discriminación o premio a este sector en el mercado de trabajo representa el 57% de los diferenciales de los ingresos por sector laboral; no obstante, cuando el modelo es corregido por el método de Heckman (1979), este porcentaje disminuye a 48%. Por otra parte, cuando se considera como estructura salarial no discriminatoria a la del sector privado, las diferencias de los salarios por sector económico se explican en un 59% (MCO) por el componente atribuido a la discriminación o penalidad a este sector laboral, sin embargo, cuando se corrige por el sesgo de selección muestral, la representatividad del componente discriminatorio en las diferencias de los salarios por sector casi no varía, pues representa el 57%.

Respecto a la parte de la brecha de ingresos explicada por las diferencias en las dotaciones grupales, los resultados muestran que si se supone como estructura salarial no discriminatoria a la del sector público, dicho componente representa un 43% y 52%, para la estimación por MCO y por Heckman, respectivamente. Finalmente, si se toma en cuenta al sector privado como estructura salarial no discriminatoria, el porcentaje que representa el componente debido a las diferencias en dotaciones en las diferencias salariales es del 41% para las estimaciones por MCO y para las estimaciones por el método de Heckman es del 43%.

---

<sup>64</sup> Donde  $\delta_{pr}$  y  $\delta_{pl}$  son parámetros específicos del grupo considerado como discriminado, para obtener mayor información al respecto véase Jann, Ben (2008) "A Stata implementation of the Blinder-Oaxaca decomposition," ETH Zurich Sociology, No 5, pp.1-24.

En resumen, existen evidencias de discriminación laboral por sector en contra del sector privado; además, los resultados del modelo no corregido, muestran que las diferencias salariales entre sectores económicos en mayor parte se debe al componente atribuido a la discriminación que a las diferencias en dotaciones; esto considerando como estructura salarial no discriminatoria tanto a la del sector público como a la del sector privado; no obstante, cuando se corrige el sesgo de selección muestral por el método planteado por Heckman(1979), estos resultados difieren de acuerdo al grupo de referencia considerado como estructura salarial no discriminatoria; entonces, cuando se toma en cuenta la estructura salarial del sector público como no discriminatoria, la diferencia salarial se explica en un mayor porcentaje por el componente atribuido a las dotaciones; pero cuando la estructura salarial del sector privado es considerada como no discriminatoria, la mayor parte de las diferencias salariales se debe al componente atribuido a la discriminación.

En relación a las diferencias salariales entre ambos sectores laborales, Carrillo (2004) expone que la diferencia es del 18% a favor de los trabajadores públicos; es decir, para el año 2003 ya existía un premio a favor de dicho sector; sin embargo, para el 2011 el premio incrementó notablemente a 67%. Así también, Panizza et al. (2005) muestran un premio de 43% a favor del sector público para el año 1995; de la misma manera, en otro de sus estudios Panizza (2000) indica que en 1998 existe un premio a favor del sector público tanto para hombres (30%) como para mujeres (26%). Cabe mencionar que las diferencias en los resultados puede deberse a los diferentes criterios por parte de cada uno de los investigadores para la toma de muestras, variables, utilización de encuestas, etc.; por ejemplo las estimaciones expuestas en Carrillo (2004) se basan en una encuesta diferente a la que este documento utilizó; pues recurrió a la Encuesta de Ingresos y Gastos de Hogares Urbanos 2002-2003 (EIGHU). Con respecto a los trabajos de Panizza (2000) y Panizza et al. (2005) no menciona una encuesta específica utilizada ya que simplemente indica que son encuestas de hogares de diferentes países latinoamericanos.



Por otro último, es importante notar que las suposiciones de tomar como estructura salarial no discriminatoria a la de los sector público o privado son útiles para entender la discriminación laboral; sin embargo, para Cahuc et al. (2004), el problema de la descomposición salarial de Blinder-Oaxaca (1973) es que la parte explicada de las diferencias salariales por sector laboral depende de cual haya sido el grupo de referencia (sector público o privado) para la descomposición salarial; ya que si se elige al sector público como grupo de referencia, los retornos de las características individuales (educación, experiencia, etc.) de este sector pueden ser mayores y además, si este grupo también está dotado con mejores características en promedio, la parte explicada de la descomposición de ingresos cuando se considera como estructura salarial no discriminatoria a la del sector público será mayor con respecto a la parte explicada de la descomposición de ingresos cuando se considera como estructura salarial no discriminatoria a la del sector privado ( $|\beta_j^{pl}(\bar{X}_j^{pr} - \bar{X}_j^{pl})| > |\beta_j^{pr}(\bar{X}_j^{pr} - \bar{X}_j^{pl})|$ ). Además, el componente atribuido a la discriminación en contra del sector privado, tomando como grupo de referencia al sector público, es más débil que tomar como grupo de referencia al otro grupo (sector privado). Para solucionar este problema Blinder-Oaxaca (1973) han propuesto un método más general que en la presente investigación no se realiza, por tanto, queda como sugerencia para posteriores trabajos.

**Tabla- 3 Descomposición de ingresos por sector**  
**“Twofold decomposition by sector”**

	Estructuras Salariales No Discriminatorias			
	Sector Público		Sector Privado	
	MCO	Heckman	MCO	Heckman
	$\beta_{pl}/(t)$	$\beta_{pl}/(t)$	$\beta_{pr}/(t)$	$\beta_{pr}/(t)$
<b>Diferencial</b>				
Predicción_1	0.681*** (66.45)	0.849*** (23.83)	0.681*** (66.45)	0.849*** (23.83)
Predicción_2	1.346*** (110.26)	1.346*** (110.26)	1.346*** (110.26)	1.346*** (110.26)
Diferencia	-0.666*** (-41.77)	-0.498*** (-13.22)	-0.666*** (-41.77)	-0.498*** (-13.22)
<b>Descomposición</b>				
Componente atribuido a Dotaciones	-0.285*** (-19.43)	-0.257*** (-18.31)	-0.273*** (-21.24)	-0.215*** (-12.45)
Componente atribuido a Discriminación	-0.380*** (-20.54)	-0.240*** (-6.23)	-0.392*** (-22.76)	-0.282*** (-10.03)

Fuente: Instituto Nacional de Estadísticas y Censo del Ecuador (INEC) - Encuesta de Empleo y Desempleo - Diciembre 2011  
 Elaboración: La autora

Nota: Las estimaciones son significativas cuando \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

(1) Entre paréntesis se encuentra el estadístico t-student.

## **CAPITULO V**

### **CONCLUSIONES Y RECOMENDACIONES**

#### **5.1. CONCLUSIONES**

Para el análisis empírico de discriminación salarial entre los sectores público y privado se emplea una metodología econométrica que en resumen se basa en tres procedimientos ya mencionados en el trabajo de investigación de Torresano (2006); (i) la estimación de ecuaciones mincerianas de ingresos controladas por sector público y privado y su interpretación mediante el análisis de sus coeficientes; (ii) estimación de un modelo multiplicativo de ingresos, igualmente controlado por sector público y privado, con interacciones para que los coeficientes varíen por sector; y (iii) en el último proceso se aplicará el método de Blinder – Oaxaca (1973), utilizando una descomposición de ingresos. Conjuntamente, en el transcurso de la investigación se aplica un modelo econométrico que permite la corrección de sesgo de selección muestral; por lo tanto, se podrá utilizar la metodología propuesta por Heckman (1979).

Consecuentemente, el análisis empírico del apartado anterior confirma la hipótesis planteada para el presente trabajo de investigación (¿Entre el Sector Público y Privado en el Ecuador para el período 2011, existe discriminación salarial?), ya que las estimaciones econométricas por el modelo multiplicativo y por el método de Blinder – Oaxaca (1973) evidencian discriminación salarial a favor del sector público y en contra del sector privado; pues las estimaciones del modelo multiplicativo corregido por el método de Heckman (1979) indican que existe un premio salarial de

aproximadamente el 28% para el sector público con respecto al sector privado; sin embargo, este modelo presenta ciertas limitaciones; como por ejemplo, supone que la estructura salarial de ambos sectores laborales son iguales, lo cual no necesariamente se cumple; por consiguiente, se podría considerar los resultados del modelo propuesto por Blinder – Oaxaca (1973) igualmente corregido por el método de Heckman (1979) que indica la existencia de un premio salarial de 50 puntos logarítmicos (o 1.51 dólares por hora) a favor del sector público en relación al sector privado; no obstante, es importante recalcar que Blinder – Oaxaca (1973) también presenta ciertas restricciones ya que el sesgo de selección muestral puede ser mayor al dividir a la muestra total en dos sub muestras; además que el número de individuos para el sector público y privado podría ser desproporcional.

En otro sentido, los resultados exponen que los asalariados tanto del sector público como privado que tienen años adicionales de experiencia, años adicionales de educación, son jefes de hogar o tienen una condición migratoria obtienen mayores ganancias. Además, ubicarse en las zonas sierra norte, sierra sur, sierra centro o Amazonía con respecto a la región costa, genera un efecto positivo en los salarios de ambos sectores laborales. En relación a la discriminación de género y etnia se encontró que en el Ecuador existe discriminación salarial en contra de las mujeres y los individuos auto identificados indígenas; particularmente, las estimaciones indican que la penalización es mayor en el sector público.

Adicionalmente, la presente investigación analiza la descomposición de las brechas salariales entre grupos planteada por Blinder-Oaxaca (1973), quienes propusieron dos componentes que contribuyen en las diferencias salariales: uno relacionado con las diferencias de las características observables de los grupos y otro atribuido a las diferencias en los rendimientos o características no observables relacionados con la discriminación salarial. Por tanto, los resultados corregidos por el método de Heckman (1979) exponen que las diferencias salariales entre ambos sectores laborales en mayor parte se deben al componente atribuido a la discriminación que a las diferencias en dotaciones; esto considerando como estructura salarial no

discriminatoria a la del sector privado; sin embargo, cuando se toma como estructura salarial no discriminatoria a la del sector público, la diferencia salarial se explica en un mayor porcentaje por el componente atribuido a las dotaciones.

Finalmente, en relación a estudios anteriores de las diferencias salariales entre ambos sectores laborales del Ecuador, para Carrillo (2004) la diferencia salarial en el año 2003 es del 18% a favor de los trabajadores públicos; también, según Panizza et al. (2005), para el año 1995, existe un premio a favor del sector público del 43%; asimismo, en otro de los estudios de Panizza (2000), para el año 1998, los hombres empleados del sector público se favorecen con un premio del 30% y las mujeres pertenecientes al mismo sector reciben un premio del 26%.

## 5.2. RECOMENDACIONES

Dado que la presente investigación concluye la existencia de diferencias salariales entre sectores laborales (público y privado) y tomando en consideración que el sector público es quien recibe una prima salarial, mientras que el sector privado recibe una penalización, convendría que el gobierno ecuatoriano formule una política en la que se obligue tanto al sector privado como al público requerir los mismos perfiles para puestos de trabajo equivalentes entre sectores; en los que los sueldos y salarios sean neutros para ambos sectores laborales.

En este mismo sentido, el Ecuador debería enfocarse en formar políticas de homologación salarial entre sectores laborales para que en el corto plazo los salarios de los sectores público y privado se equiparen; considerando similares características entre los trabajadores públicos y privados. Además, estas políticas ayudarían en la disminución del desempleo porque como se mencionó al inicio de este trabajo: “En el caso de que se encuentren diferencias salariales, no necesariamente esto significa que la eficiencia del mercado laboral esté distorsionada, ya que para Carrillo (2004) basándose en la teoría económica convencional se manifiesta que si el sector privado ecuatoriano es competitivo, entonces el salario fijado por este, sería eficiente y equivalente a la productividad laboral de sus trabajadores, pero si el sector público fija sueldos más altos, el equilibrio del mercado laboral se distorsionaría y como consecuencia el desempleo se incrementaría. En otro sentido, si el sector privado en el país tiene poder en el mercado, entonces los salarios fijados por este sector serían inferiores al de eficiencia y con salarios mejor remunerados por parte de la administración pública, existiría presión para eliminar las distorsiones, incrementando el bienestar de la sociedad.”

Por otra parte, el gobierno ecuatoriano podría promover incentivos al sector privado como las reducciones de impuestos con la condición de que incrementen los salarios

de los trabajadores con características equivalentes a trabajadores públicos que adquieren un ingreso mayor.

También, se recomienda analizar indicadores macroeconómicos para resolver el problema de discriminación salarial. Por ejemplo, bajo el supuesto de que el sector privado no presenta buenos rendimientos en comparación con el sector público; entonces, su productividad laboral debería mejorar con el objetivo de que sus ingresos aumenten.

Adicionalmente, los resultados indican que en cuanto al género y etnia existe discriminación entre sectores laborales; por ende, se podría sugerir aplicar políticas en las que se promocióne la igualdad de género y etnia.

Entre las recomendaciones metodológicas se pueden citar las más relevantes como realizar en futuras investigaciones el método de regresión de cuantiles que propone Carrillo (2004), para identificar las diferencias salariales en cada cuantil de la distribución salarial; además, se podría combinar las diferencias salariales entre los sectores laborales, el género y la etnia.

Para corregir el sesgo de endogeneidad de la educación se podrían utilizar variables instrumentales que en la presente investigación no se realiza por falta de información. Adicionalmente, se recomendaría que para conocer la denominada pre-discriminación salarial se debería recolectar información acerca del historial familiar<sup>65</sup>.

Finalmente, en el presente trabajo solamente se efectuó la descomposición en dos componentes denominada "twofold descomposition", cuya limitación consiste en que se debe tomar como estructura salarial no discriminatoria a la de un grupo de referencia; por tanto, para solucionar dicho problema, se podría aplicar para futuras investigaciones el modelo más general propuesto por Blinder-Oaxaca (1973), en el que no se considera a alguna estructura salarial en especial como no discriminatoria.

---

<sup>65</sup> Considera si el padre o la madre se ha educado, los ingresos de los padres, etc.

## BIBLIOGRAFÍA

- Albert, Cecilia y Moreno, Gloria (1998), “*Diferencias salariales entre el sector público y privado español: un modelo de switching*”, Estadística Española, Vol. 40, Núm. 143, 1998, pp.167 a 193.
- Altonji, Joseph y Blank, Rebecca (1999), “*Race and gender in the labor market*”, en O. Ashenfelter y D. Card, Handbook of Labor Economics, Elsevier Science, pp. 3143-3258.
- Banco Central del Ecuador (BCE), página web: [www.bce.fin.ec](http://www.bce.fin.ec), último acceso: 25 de febrero de 2013.
- Becker, Gary (1973), “*The economics of discrimination*”, The University of Chicago press, 2da edición, Chicago.
- Belsley, D. A., E. Kuh, and R. E. Welsch (1980), “*Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*”. New York: Wiley, pp. 1-321.
- Blinder, Alan (1972), “*Wage discrimination: reduced form and structural estimates*”, The Journal of Human resources, Vol. 4, pp. 437-455.
- Breusch, T. S., and A. R. Pagan. (1979), “*A simple test for heteroscedasticity and random coefficient variation*”, The Journal of Econometric Society, Econométrica 47, pp. 1287–1294.
- Cain, Glen (1986), “*The economic analysis of labor market discrimination*”, en O. Ashenfelter y R. Layard Handbook of Labor Economics, Elsevier Science, pp. 693-695.



- Cahuc P. y Zylberberg A. (2004), *“Labor Economics”*, MIT Press
- Carrillo, Paúl (2004), *“Las diferencias salariales entre el sector público y privado en el Ecuador”*, Cuestiones Económicas (Quito), Volumen 20, N. 2, pp. 165-173.
- Carrillo, Paúl y Vásconez, Verónica (2010), *“An Approximation of the Wage Gap between Public Entities in Ecuador”*, Munich Personal RePEc Archive.
- Casto, Alfonso (2008), *“Econometría: Modelos Estáticos”*, Cuadernos de Matemática, Escuela Politécnica Nacional, pp. 43-151.
- Casto, Alfonso (2010), *“Regresión lineal”*, Monografías de Matemática y Estadística, Escuela Politécnica Nacional, pp. 1-192.
- Código de Trabajo Ecuatoriano, capítulo V, artículo 47 y artículo 55.
- Cook, R. D. (1977), *“Detection of influential observation in linear regression”*, Technome-Trics 19, pp. 15–18.
- De Holanda, Ana y De Holanda, Fernando (2012), *“Diferencial de salarios entre os setores público e privado no Brasil: um modelo de escolha endógena”*, Instituto de Investigación de economía Aplicada, Texto para el Debate No. 1713, Río de Janeiro.
- “Encuesta de empleo, desempleo y subempleo”. Instituto Nacional de Estadísticas y Censos (INEC), 2011.
- Galarza, Pablo (2009), *“Inclusión de la variable etnia en las fuentes de información sociodemográfica del Ecuador”*, CEPAL, Santiago de Chile.
- Galindo, Edwin (2006), *“Estadística métodos y aplicaciones para administración e ingeniería”*, ProCiencia Editores, Quito.
- Harmon, Colm et al. (2003), *“The returns to education: microeconomics”*, Journal of economic surveys, Vol. 17, No. 2, pp. 115-155.

- Heckman, James (1979), "*Sample Selection Bias as a Specification Error*", *Econometría*, Vol. 47, No. 1, pp. 153-161
- Heitmueller, Axel (2006), "*Public-Private Sector Pay Differentials In A Devolved Scotland*", *Journal of Applied Economics*, Vol IX, No. 2, pp. 295-323.
- Iglesias Garrido, Jesús (2005), "*Capital Humano y Señalización*", Universidad Autónoma de Barcelona, pp. 5-8.
- Instituto Nacional de Estadísticas y Censos (INEC). página web: [www.inec.gob.ec](http://www.inec.gob.ec), último acceso: 25 de febrero de 2013.
- Jann, Ben (2008) "*A Stata implementation of the Blinder-Oaxaca decomposition*", *ETH Zurich Sociology*, No 5, pp.1-24.
- López, Erique y Motellón Elizabet (2008), "*Descomposición de diferencias salariales regionales en presencia de selección muestral*", Universidad de Barcelona y Universidad Oberta de Catalunya.
- Mejía, José y Moncada, Gilberto (2000), "*Las variables de etnia y raza en las encuestas de hogares en América Latina y el Caribe*", Ponencia presentada en el Primer Encuentro Internacional "Todos Contamos. Los Grupos Étnicos en los Censos", Cartagena.
- Mincer, J. (1974), "*Schooling, Experience and Earnings*", Columbia University Press New York.
- Ministerio de Finanzas, página web: [www.finanzas.gob.ec](http://www.finanzas.gob.ec), acceso: 25 de febrero de 2013.
- Neuman, Shoshana y Oaxaca, Ronald (2004), "*Wage decompositions with selectivity-corrected wage equations: A methodological note*", *Journal of Economic Inequality*, pp. 3-10.

- Oaxaca, Ronald (1973), "*Male-Female wage differentials in urban labor markets*", *International Economic Review*, Vol. 14, No. 3, pp. 693-709.
- Organización Internacional del Trabajo (OIT), página web: [www.ilo.org](http://www.ilo.org), acceso: 16 de noviembre de 2012.
- Panizza, Ugo (2000), "*The public sector premium and the gender gap in Latin America: evidence for the 1980s and 1990s*", Inter-American Development Bank, Working Paper No. 431, pp. 1-25.
- Panizza, Ugo y Zhen-Wei Qiang, Christine (2005), "*Public-private wage differential and gender gap in Latin America: spoiled bureaucrats and exploited women?*", *Journal of Socio-Economics*, Vol. 34, pp. 810-833.
- Rama, Martín (2003), "*The Sri Lankan Unemployment Problem Revisited*", Banco Mundial, *Review of Development Economics*, pp. 510–525.
- Torresano, Daniel (2009), "*Análisis empírico de la discriminación salarial por género y etnia en el Ecuador en el año 2008*", Universidad Autónoma de Barcelona, Tesina.
- Willis, Robert (1986), "*Wage determinants: A survey and reinterpretation of human capital earning functions*", O. Ashenfelter y R. Layard *Handbook of Labor Economics*, Elsevier Science, pp. 525-602.

## ANEXOS

### ANEXO- A Validación del modelo multiplicativo

Para encontrar el mejor modelo multiplicativo que se ajuste a los datos y que permita entender la realidad respecto a las diferencias salariales por sector laboral, se realizó pruebas con varias interacciones de variables. Un primer modelo corregido por el método de Heckman (1979) se muestra a continuación:

Source	SS	df	MS			
Model	932.098211	20	46.6049105	Number of obs =	5542	
Residual	1438.78423	5521	.260602107	F( 20, 5521) =	178.84	
Total	2370.88244	5541	.427879885	Prob > F =	0.0000	
				R-squared =	0.3931	
				Adj R-squared =	0.3909	
				Root MSE =	.51049	

w	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
educacion	.0590469	.0039354	15.00	0.000	.0513319	.0667618
experiencia	.0150907	.0020119	7.50	0.000	.0111467	.0190348
experiencia2	-.0001036	.0000399	-2.60	0.009	-.0001818	-.0000255
publico	.2091991	.071387	2.93	0.003	.0692525	.3491458
mujer	-.1075371	.0209513	-5.13	0.000	-.1486099	-.0664642
indigena	-.131234	.0674131	-1.95	0.052	-.2633902	.0009221
jefe_hogar	.0748678	.0272342	2.75	0.006	.021478	.1282577
migracion	.0614171	.0197919	3.10	0.002	.0226172	.100217
nor_sierra	.0718351	.019832	3.62	0.000	.0329566	.1107136
ctro_sierra	.0702489	.0206597	3.40	0.001	.0297477	.1107501
sur_sierra	.0989799	.0219427	4.51	0.000	.0559636	.1419962
amazonia	.1851301	.0344839	5.37	0.000	.1175281	.252732
guayaquil	.140944	.0290383	4.85	0.000	.0840175	.1978705
publico_educacion	.0056845	.0039962	1.42	0.155	-.0021497	.0135186
publico_expe	.0043493	.0013235	3.29	0.001	.0017548	.0069438
publico_mujer	-.0325295	.0321874	-1.01	0.312	-.0956294	.0305704
publico_ind	.0114573	.0866045	0.13	0.895	-.1583215	.1812361
publico_jefe	-.0523316	.0357283	-1.46	0.143	-.1223731	.0177099
publico_gye	-.04423	.0551758	-0.80	0.423	-.1523963	.0639363
invmills2	-.2139495	.0724758	-2.95	0.003	-.3560307	-.0718683
_cons	-.1560813	.0889792	-1.75	0.079	-.3305155	.0183529

Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

Dado que el presente modelo contiene interacciones no significativas se comprueba mediante la siguiente prueba de hipótesis si el modelo multiplicativo debe o no estar conformado por las variables de interacción.

**Prueba de Hipótesis:**

*Ho: público X educación = público X experiencia = público X mujer = público X indígena = público X jefe\_hogar = público X guayaquil = 0*

*H1: público X educación ≠ 0 o público X experiencia ≠ 0 o público X mujer ≠ 0 o público X indígena ≠ 0 o público X jefe\_hogar ≠ 0 o público X guayaquil ≠ 0*

chi2(6)	11.47
Prob> chi2	0.0749

Se acepta la hipótesis nula al nivel 5% pero se rechaza al nivel 8%; por consiguiente, se verifica las hipótesis individuales observando el estadístico t y se eliminan las variables no significativas. Por tanto, el nuevo modelo es así:

Source	SS	df	MS	Number of obs = 5542		
Model	930.942546	15	62.0628364	F( 15, 5526) =	238.18	
Residual	1439.9399	5526	.260575443	Prob > F =	0.0000	
Total	2370.88244	5541	.427879885	R-squared =	0.3927	
				Adj R-squared =	0.3910	
				Root MSE =	.51047	

w	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
educacion	.0614625	.0032422	18.96	0.000	.0551065	.0678185
experiencia	.0160188	.0019257	8.32	0.000	.0122438	.0197939
experiencia2	-.0001081	.0000396	-2.73	0.006	-.0001858	-.0000305
publico	.2767046	.0253569	10.91	0.000	.226995	.3264142
mujer	-.120457	.0159468	-7.55	0.000	-.1517189	-.0891951
indigena	-.1243873	.0429773	-2.89	0.004	-.2086396	-.040135
jefe_hogar	.0520416	.0204641	2.54	0.011	.0119239	.0921592
migracion	.0619626	.0197563	3.14	0.002	.0232325	.1006928
nor_sierra	.0714392	.0197676	3.61	0.000	.032687	.1101915
ctro_sierra	.070484	.0206178	3.42	0.001	.0300649	.110903
sur_sierra	.0988061	.0219183	4.51	0.000	.0558375	.1417746
amazonia	.1855015	.0343507	5.40	0.000	.1181606	.2528424
guayaquil	.127831	.0257706	4.96	0.000	.0773106	.1783515
publico_expe	.0029303	.0011033	2.66	0.008	.0007673	.0050932
invmills2	-.2107907	.0694636	-3.03	0.002	-.3469666	-.0746147
_cons	-.1855233	.0778739	-2.38	0.017	-.3381869	-.0328598

A continuación se prueba nuevamente si el modelo debe o no contener la interacción entre las variables público y experiencia.

$H_0$ : público  $\times$  experiencia = 0

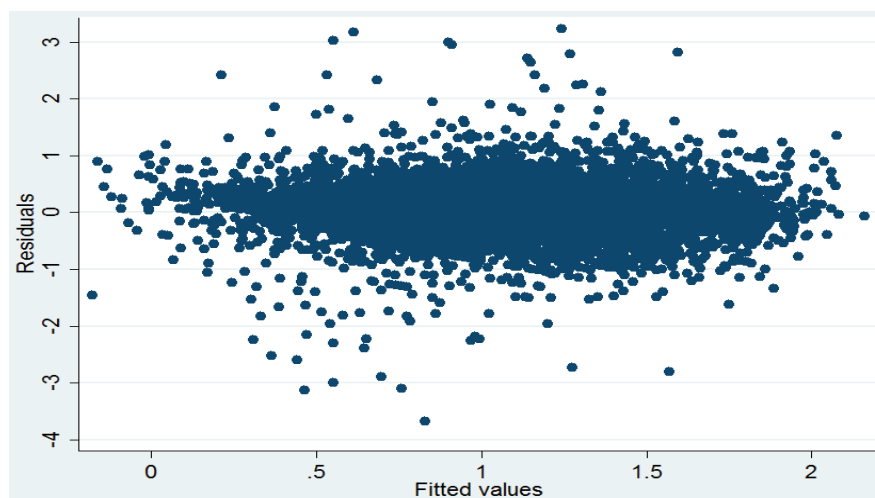
$H_1$ : público  $\times$  experiencia  $\neq$  0

chi2(1)	7.08
Prob> chi2	0.0078

Se rechaza la hipótesis nula al nivel 1%. En conclusión no se debe omitir la interacción entre las variables público y experiencia.

Antes de analizar los resultados del último modelo multiplicativo, se verifica si este cumple con las hipótesis planteadas<sup>66</sup>. A continuación se prueba la linealidad del modelo mediante gráficos de residuos vs el pronóstico de la variable dependiente y las variables explicativas; así se tiene:

**Gráfico de residuos en función del pronóstico**



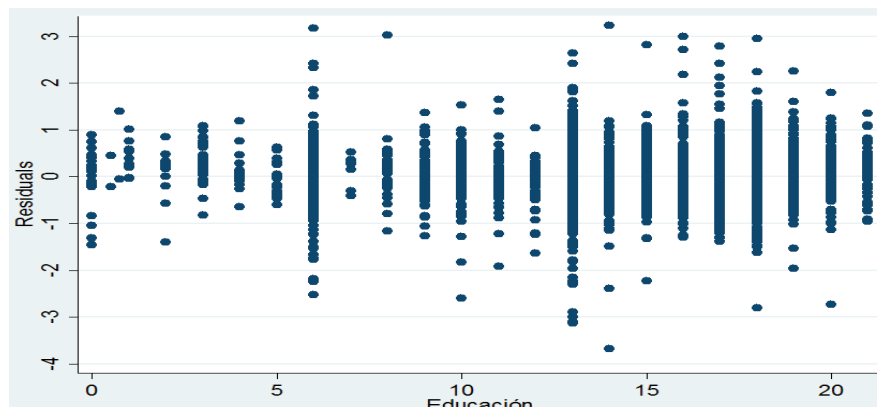
Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

Este gráfico muestra una nube de puntos que sugiere que la relación es lineal; además, no muestra evidencias de heterocedasticidad ya que se observa que las observaciones están distribuidas al azar en una franja horizontal simétrica; sin

<sup>66</sup> Para una mejor comprensión con respecto a la validación del modelo véase Castro (2008), pp. 33-52 y 112-115.

embargo, existen pocos puntos que se encuentran fuera del intervalo  $[-3, 3]$ , dando la sospecha de la existencia de puntos atípicos.

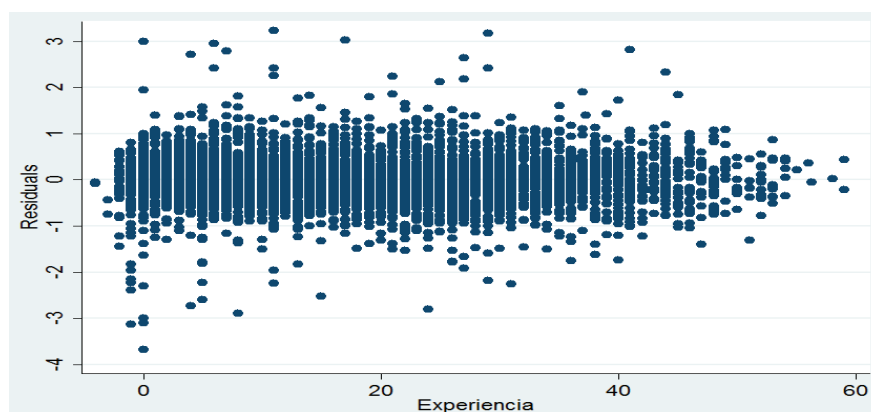
**Gráfico de residuos en función de la educación**



Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

Las observaciones en este gráfico muestran ciertos puntos atípicos; pero no muestra problemas con la varianza y la hipótesis de linealidad. Por consiguiente, se podría sospechar que el modelo es válido para entender las diferencias salariales entre los sectores público y privado.

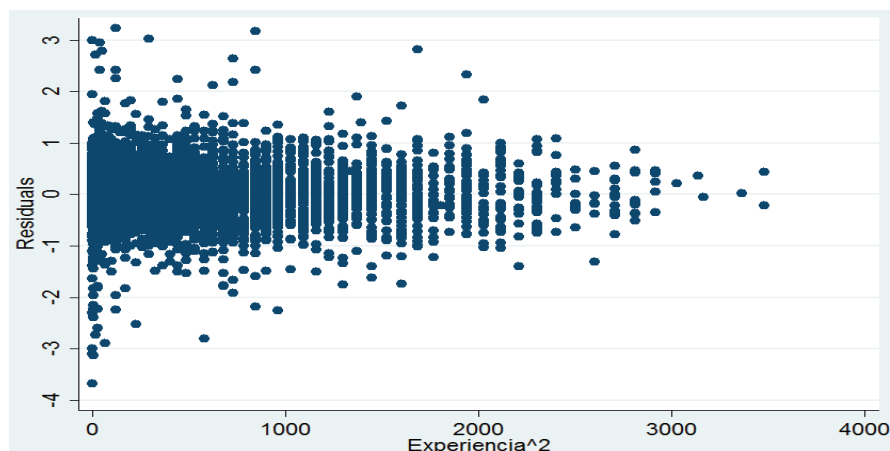
**Gráfico de residuos en función de la experiencia**



Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

Al igual que los gráficos anteriores, este no refleja violación de las hipótesis de linealidad, homocedasticidad, normalidad y correlación; no obstante, se observa la presencia de puntos atípicos.

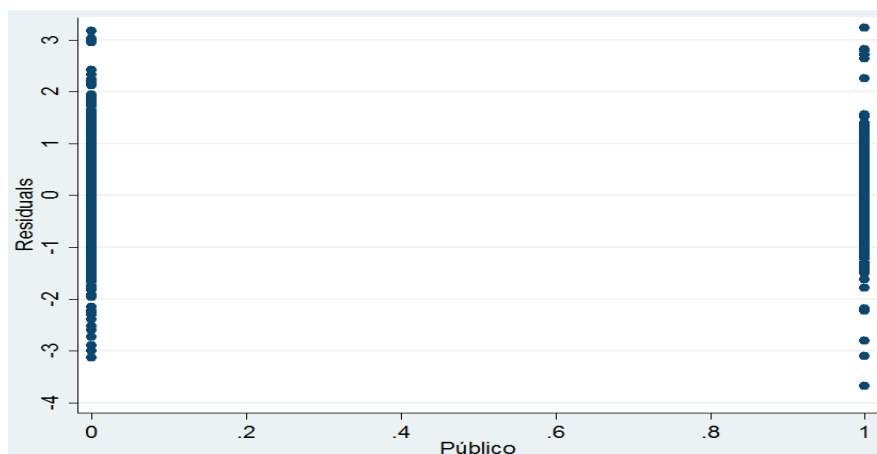
### Gráfico de residuos en función de la experiencia al cuadrado



Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

El gráfico de residuos en función de la experiencia al cuadrado muestra evidencias de que la varianza no es constante; por tanto, esta hipótesis se verificará más adelante. Adicionalmente, en el gráfico existen 4 puntos atípicos que deberán ser estudiados.

### Gráfico de residuos en función de sector laboral

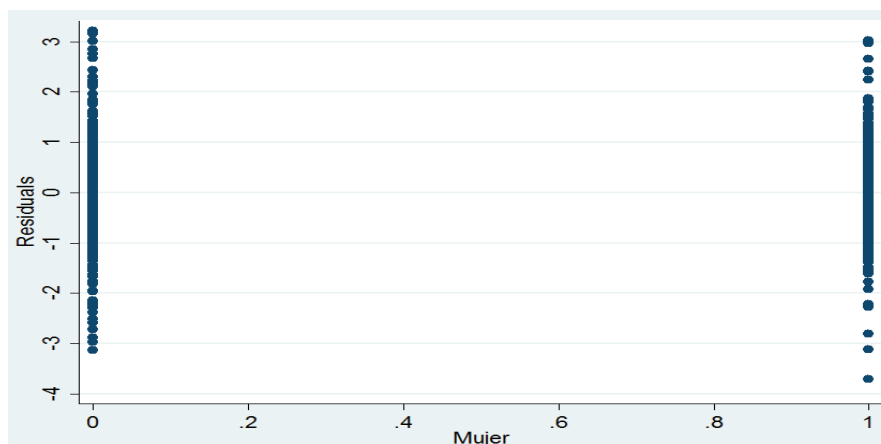


Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

En el presente gráfico la variable público es dicotómica que toma el valor de 1 si el individuo pertenece al sector público y 0 si pertenece al sector privado, es por esto que los puntos forman dos líneas verticales. Entonces, el gráfico no evidencia violación de las hipótesis pero sí muestra puntos atípicos.



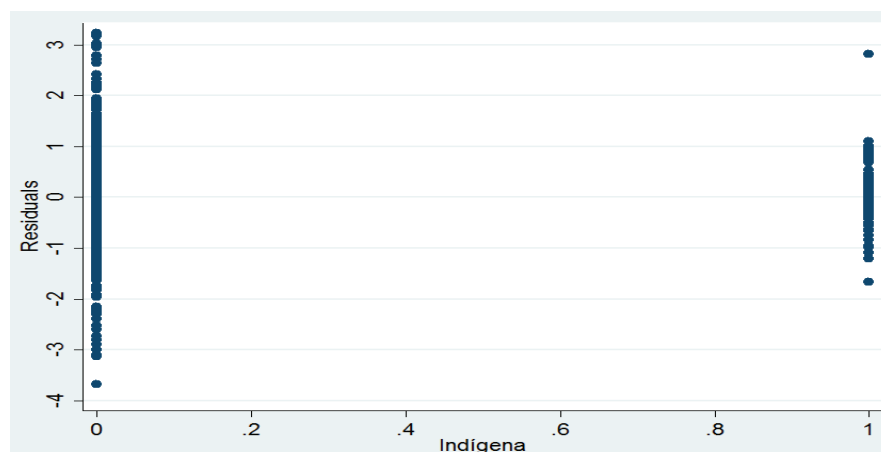
### Gráfico de residuos en función de género



Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

Paralelamente al gráfico anterior la variable mujer es dicotómica (toma el valor de uno si el individuo es mujer y 0 si es hombre) y en la formación de las dos líneas verticales solamente se observa la existencia de ciertos puntos atípicos.

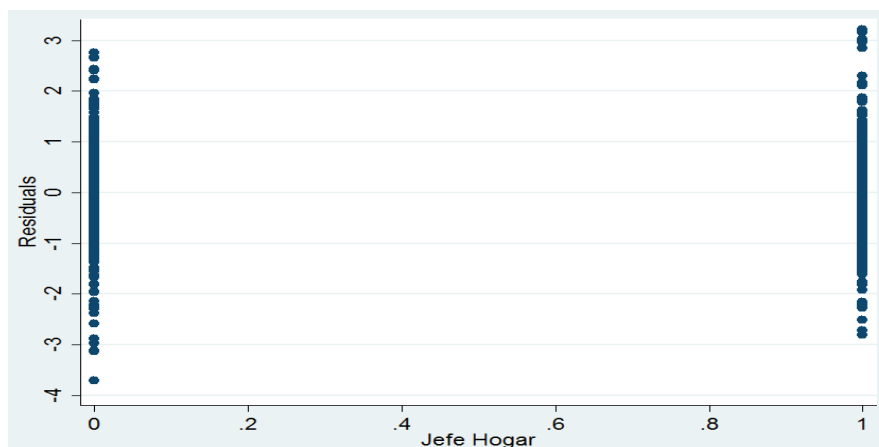
### Gráfico de residuos en función de auto identificación étnica



Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

En el gráfico de la variable indicatriz indígena se identifica 3 puntos atípicos solamente en el grupo de individuos que no se consideran como parte de la población indígena.

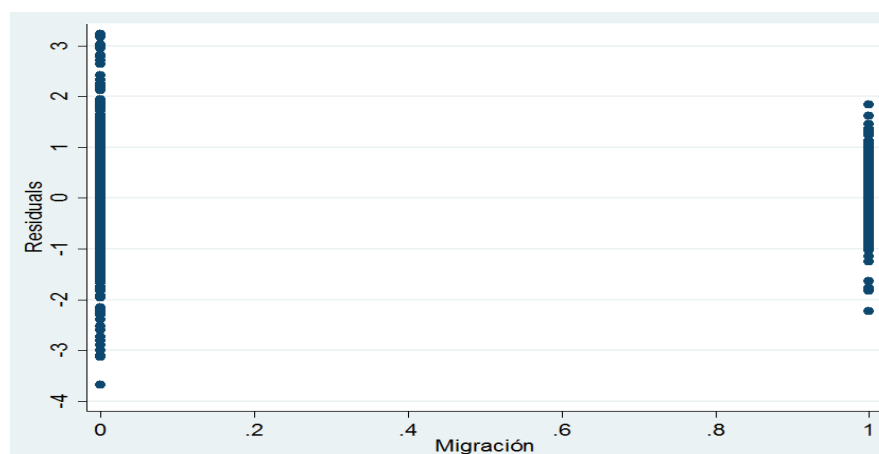
### Gráfico de residuos en función de relación de parentesco



Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

Como es evidente, el gráfico de residuos frente la variable jefe de hogar muestra dos grupos de individuos, el primer grupo representa a las personas que no son jefes de hogar y el segundo considera a quienes se identifican como jefes de hogar. En ambos grupos existen puntos atípicos que son estudiados más adelante.

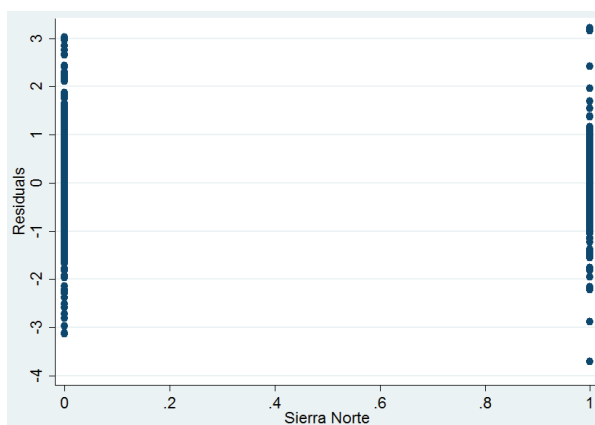
### Gráfico de residuos en función de la situación migratoria



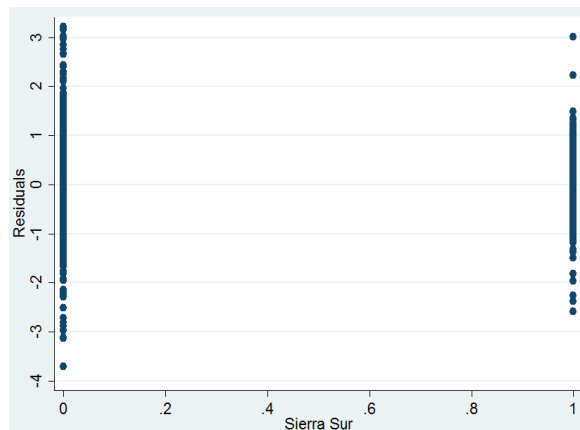
Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

En las dos poblaciones que muestra el gráfico precedente no se evidencia violación de las hipótesis; sin embargo, solamente en el grupo conformado por individuos que no han migrado se identifica 4 puntos atípicos.

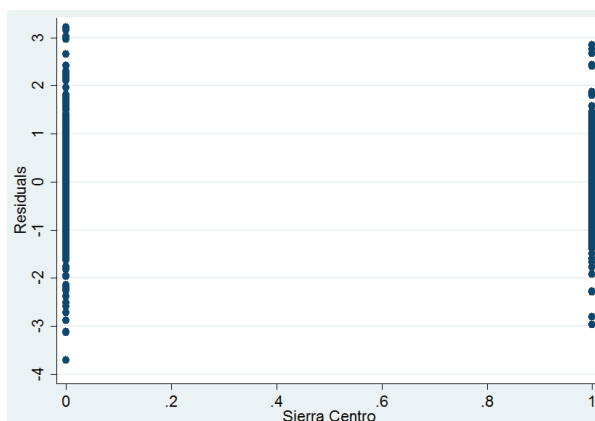
**Gráfico de residuos en función de la ubicación sierra norte**



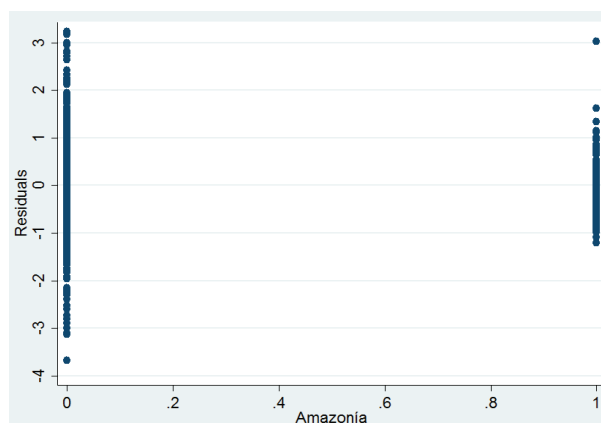
**Gráfico de residuos en función de la ubicación sierra sur**



**Gráfico de residuos en función de la ubicación sierra centro**



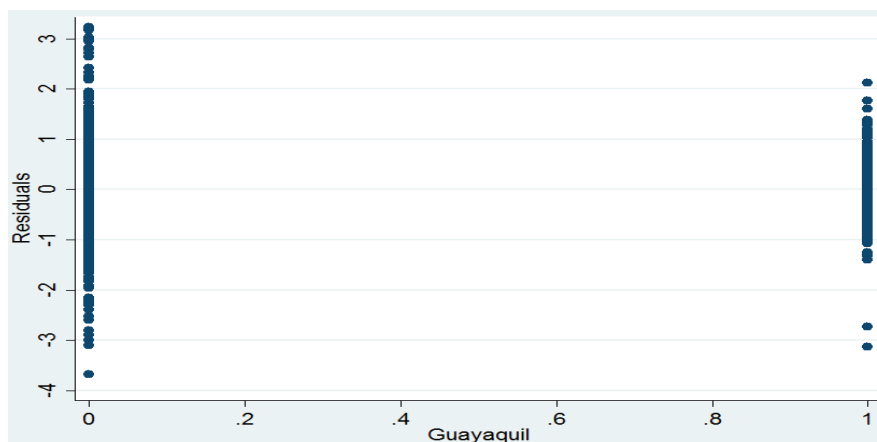
**Gráfico de residuos en función de la región amazónica**



Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

Estos gráficos representan a la variable de localización geográfica y muestran que no hay indicios para que las hipótesis no se cumplan a pesar de que cada gráfico indica la existencia de ciertos puntos atípicos. Además, como era de esperarse cada gráfico forma dos poblaciones, pues las variables son binarias.

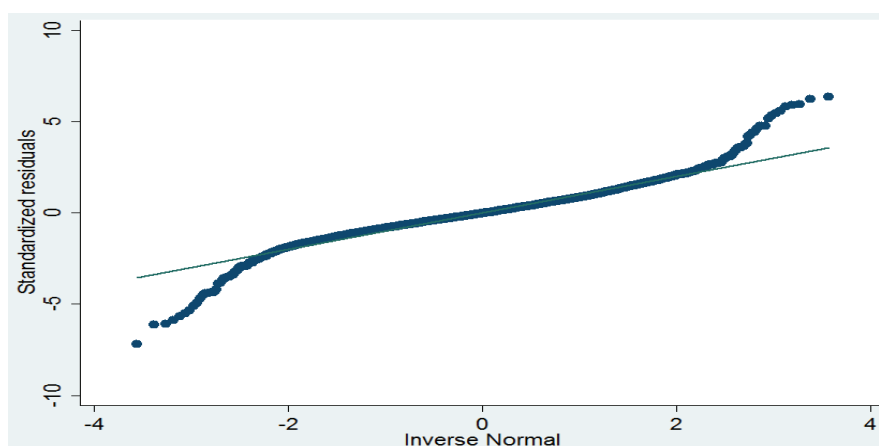
### Gráfico de residuos en función de la ubicación ciudad de Guayaquil



Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

El presente gráfico está conformado por dos líneas verticales que representan a dos poblaciones diferentes. A la primera población pertenecen los individuos que no viven en la ciudad de Guayaquil, aquí se evidencia puntos atípicos pero no se sospecha el incumplimiento de las hipótesis; lo mismo ocurre con el segundo grupo poblacional conformado por individuos que habitan en Guayaquil.

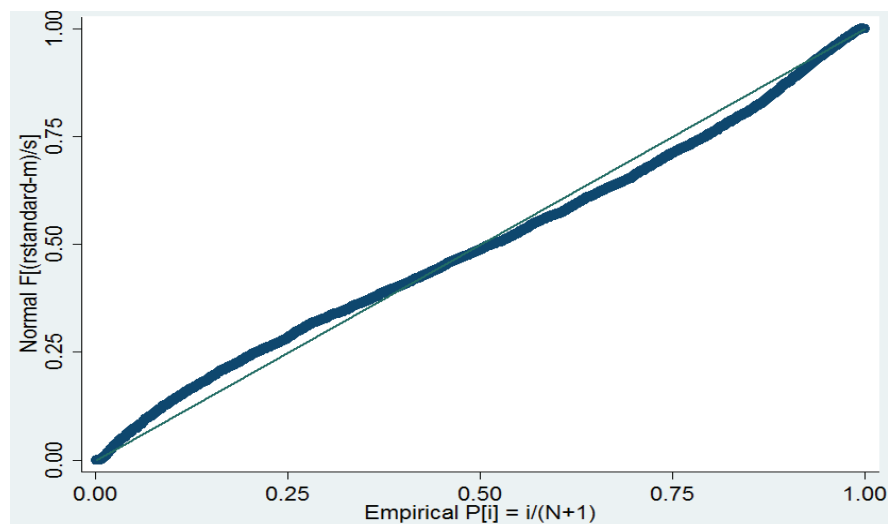
### Gráfico de cuantiles de los residuos estandarizados frente a los cuantiles de la distribución normal



Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

El gráfico de cuantiles indica la violación de la hipótesis de normalidad de los residuos ya que se observa tanto en el extremo derecho como en el izquierdo puntos que se alejan de la recta.

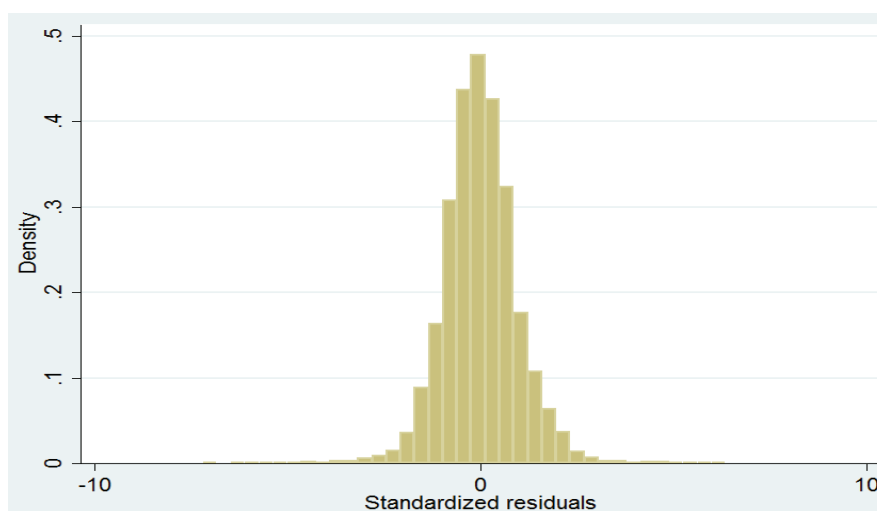
### Gráfico de percentiles de los residuos estandarizados frente a los percentiles de la distribución normal



Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

También, el gráfico de percentiles muestra evidencia de la violación de la hipótesis de normalidad ya que no todos los puntos se ajustan a la línea recta.

### Histograma de residuos estandarizados



Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

Aparentemente el histograma muestra que los residuos se aproximan a una ley normal; no obstante, se observa apuntamiento y asimetría.

### Prueba de apuntamiento y asimetría de los residuos estandarizados

Variable	Pr(Asimetría)	Pr(Curtosis)	Prob>chi2
rstandard	0.0357	0.0000	0.0000

Fuente: ENEMDU, diciembre 2011

Elaboración: La autora

En conclusión, de acuerdo a la interpretación que Carrillo (2008) plantea, en el histograma y en el gráfico de probabilidad normal se observa que un 95% de los residuos estandarizados se encuentran entre -2 y 2, y el 99% entre -3 y 3; sin embargo, en las pruebas de normalidad se acepta la hipótesis nula de existencia de simetría ( $H_0: CA=0$ ) y se rechaza la hipótesis nula de que el coeficiente de apuntamiento (CAp) sea igual a 3; por tanto, dado que solo la prueba de hipótesis de apuntamiento es significativa, se sospecha de que los errores siguen una ley diferente a la normal (se rechaza hipótesis nula de normalidad de los residuos); por consiguiente, se verifica la existencia de datos atípicos ya que en el primer gráfico de probabilidad normal existen puntos ubicados al inicio y al final del gráfico que se alejan de la recta. La eliminación de estos puntos probablemente corrijan la normalidad de los residuos.

Para identificar a los puntos palanca se obtiene los elementos de la diagonal de la matriz de proyección denominada “matriz sombrero”. Para Castro (2008) si estos valores son mayores al límite  $27/n$ , entonces el punto correspondiente es un punto palanca. Así también, para reconocer a los puntos influyentes se obtiene la distancia de Cook, cuyos valores mayores a  $4/n$  son considerados puntos influyentes<sup>67</sup>. El paquete estadístico Stata (2012) obtiene la distancia de Cook así:

$$D_j = \frac{\hat{e}_{sj}^2 (s_{pj} s_{rj})^2}{k} = \frac{h_j \hat{e}_j^2}{k s^2 (1 - h_j)^2}$$

<sup>67</sup> Al respecto véase Stata base reference manual, release 12, Stata Corp LP, College Station, Texas, pp. 1731-1733.

Donde,  $k$  es el número de regresores más la constante,  $s_{pj}$  es el error estándar de la predicción,  $s_{rj}$  el error estándar de los residuos y  $\hat{e}_{sj} = \hat{e}_j/s_{rj}$  son los residuos estandarizados.

Por otra parte, para los puntos atípicos se calculan los residuos estudentizados; así el paquete estadístico Stata (2012) los calcula mediante la siguiente fórmula:

$$r_j = \frac{\hat{e}_j}{s_{(j)}\sqrt{1-h_j}}$$

Sea  $\hat{e}_j$  los residuos,  $s_{(j)}$  representa la raíz del error cuadrático medio eliminando la  $j$ th observación y  $h_j$  es el elemento de la diagonal de la matriz de proyección “sombrero”. Por tanto, según Castro (2008), si  $r_j$  es mayor a 3 se considera punto atípico.

El resumen estadístico de los puntos Palanca, Influyentes o Atípicos están descritos en las siguientes tablas:

#### Puntos palanca

Variable	Obs	Mean	Std. Dev.	Min	Max
ppalanca	477	.0072686	.0022142	.0048731	.0152882

Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

#### Puntos influyentes

Variable	Obs	Mean	Std. Dev.	Min	Max
pinfluye	272	.0018005	.0018256	.0007222	.018104

Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

#### Puntos atípicos

Variable	Obs	Mean	Std. Dev.	Min	Max
patipico	17	5.173892	.7193994	4.164289	6.371505

Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

### Puntos palanca influyentes no atípicos

Variable	Obs	Mean	Std. Dev.	Min	Max
ppalanca	69	.0077971	.0024321	.0048968	.0150531

Fuente: ENEMDU, diciembre 2011

Elaboración: La autora

### Puntos palanca influyentes atípicos

variable	Obs	Mean	Std. Dev.	Min	Max
ppalanca	4	.0068088	.0016834	.0057702	.0093248

Fuente: ENEMDU, diciembre 2011

Elaboración: La autora

Se eliminarán de la muestra las observaciones que se muestran a continuación:

### Puntos palanca influyentes y atípicos

Observación	$h_{ii}$	$D_j$	$r_j$
484.	.0093248	.018104	5.562433
4253.	.006048	.0134416	5.963729
5376.	.0060921	.0050146	3.621982
5487.	.0057702	.003634	3.167787

Fuente: ENEMDU, diciembre 2011

Elaboración: La autora

Con la nueva muestra se obtiene el siguiente modelo estimado por el método de Heckman (1979):



Source	SS	df	MS			
Model	930.742712	15	62.0495141	Number of obs =	5538	
Residual	1416.68555	5522	.256552979	F( 15, 5522) =	241.86	
Total	2347.42826	5537	.42395309	Prob > F =	0.0000	
				R-squared =	0.3965	
				Adj R-squared =	0.3949	
				Root MSE =	.50651	

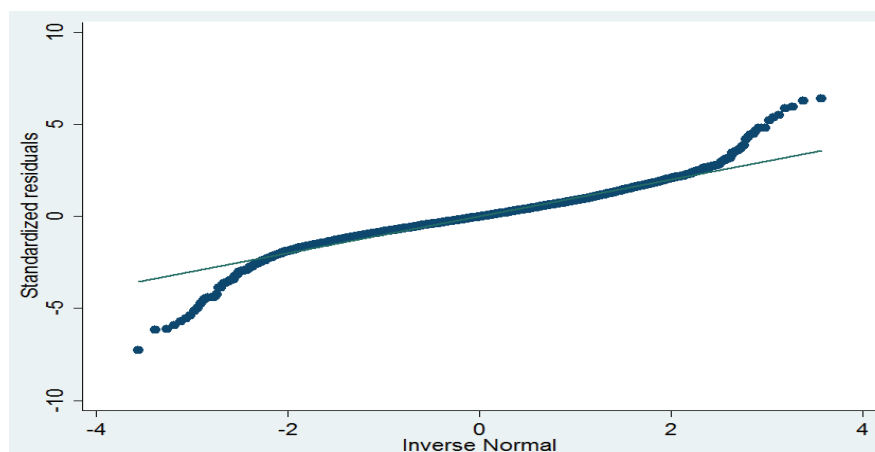
  

w	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
educacion	.0610244	.0032175	18.97	0.000	.0547169	.067332
experiencia	.0160489	.0019112	8.40	0.000	.0123021	.0197957
experiencia2	-.0001125	.0000393	-2.86	0.004	-.0001896	-.0000353
publico	.2802052	.0251679	11.13	0.000	.2308662	.3295442
mujer	-.1222162	.0158327	-7.72	0.000	-.1532546	-.0911779
indigena	-.1385937	.0427941	-3.24	0.001	-.2224869	-.0547004
jefe_hogar	.0480615	.0203002	2.37	0.018	.0082652	.0878578
migracion	.0594957	.0196269	3.03	0.002	.0210192	.0979721
nor_sierra	.0714039	.0196148	3.64	0.000	.0329511	.1098567
ctro_sierra	.0655599	.0204687	3.20	0.001	.0254331	.1056867
sur_sierra	.099049	.0217489	4.55	0.000	.0564126	.1416853
amazonia	.1696928	.0342041	4.96	0.000	.1026393	.2367463
guayaquil	.1280423	.0255712	5.01	0.000	.0779127	.1781718
publico_expe	.0029024	.0010955	2.65	0.008	.0007547	.00505
inv Mills3	-.2223612	.0688744	-3.23	0.001	-.3573821	-.0873404
_cons	-.1727497	.0772491	-2.24	0.025	-.3241885	-.021311

Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

Nuevamente se valida el nuevo modelo así:

### Gráfico de cuantiles de los residuos estandarizados frente a los cuantiles de la distribución normal

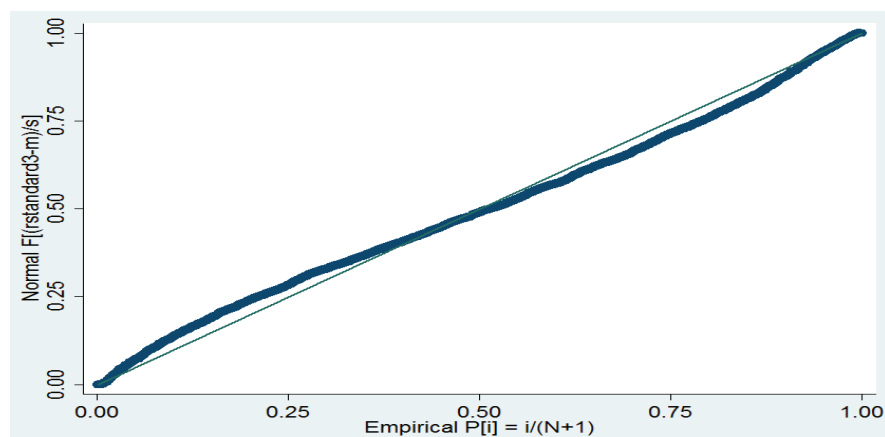


Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

El gráfico de cuantiles de la nueva regresión muestra que no se ha podido corregir la normalidad con la eliminación de los puntos palanca, influyentes y atípicos; ya que

las observaciones del inicio y final del gráfico se alejan de la recta de la distribución normal.

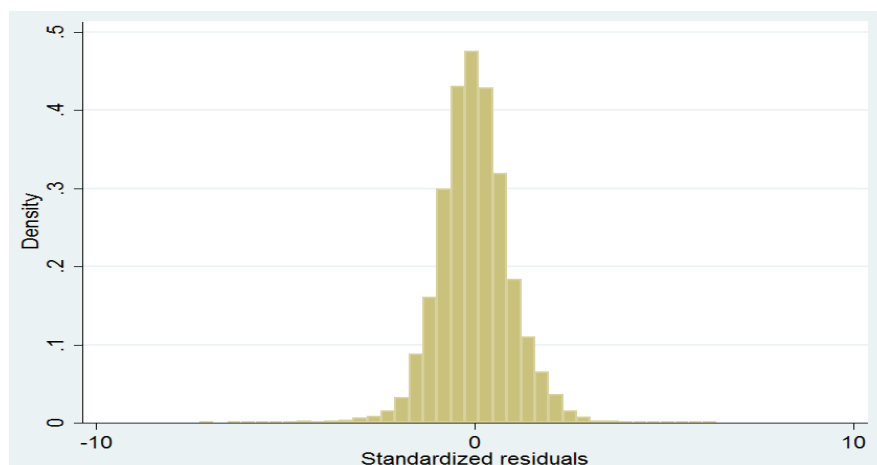
### Gráfico de percentiles de los residuos estandarizados frente a los percentiles de la distribución normal



Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

Igualmente, no se observa un cambio en el nuevo gráfico de percentiles sugiriendo que con la eliminación de los puntos mencionados no se podrá corregir la normalidad de los residuos.

### Histograma de residuos estandarizados



Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

La asimetría parece mejorar con respecto al anterior histograma; aunque el apuntamiento parezca ser el mismo.

### Prueba de apuntamiento y asimetría de los residuos estandarizados

Variable	Pr(Asimetría)	Pr(Curtosis)	Prob>chi2
rstandard	0.9010	0.0000	0.0000

Fuente: ENEMDU, diciembre 2011

Elaboración: La autora

En síntesis, los gráficos anteriores son indicadores de normalidad que al parecer no se han corregido con respecto a los gráficos de la primera regresión; además, existen ciertos puntos atípicos y palanca que se mantienen. Por otra parte, haciendo referencia a la interpretación que Carrillo (2009) expone, se dice que en las pruebas de normalidad se acepta la hipótesis nula de existencia de simetría ( $H_0: CA=0$ ) y se rechaza la hipótesis de que el coeficiente de apuntamiento (CAp) sea igual a 3; por consiguiente, existe la sospecha de que los errores siguen una ley diferente a la normal; es decir se rechaza la hipótesis nula de normalidad de los residuos; en consecuencia, se prefiere el primer modelo aunque este posea puntos atípicos, puntos palanca e influyentes; pues la causa puede ser variables no incluidas en el modelo como por ejemplo regresores que representan el historial familiar (educación/ocupación/ingreso del padre, número de hermanos, historia migratoria, etnicidad, etc.), puesto que no hay disponibilidad de información al respecto.

Dado que se prefiere al primer modelo multiplicativo con respecto al segundo, se debe continuar con el análisis de los errores; de ahí que se realizan pruebas de multicolinealidad y heterocedasticidad.

#### Prueba de multicolinealidad

Para detectar problemas de Multicolinealidad del modelo, el paquete estadístico Stata (2012) calcula el Factor Inflación de la Varianza centrado (VIFc<sup>68</sup>) a través de la siguiente fórmula:

$$VIFc(X_j) = \frac{1}{1 - \hat{R}_j^2}$$

---

<sup>68</sup> Por sus siglas en inglés.

Donde  $\hat{R}_j^2$  es el cuadrado del coeficiente de correlación múltiple centrado en la regresión de  $X_j$  en función de los restantes  $k-2$  regresores.

A continuación se muestra el VIF de cada regresor:

Variable	VIF	1/VIF
experiencia	13.59	0.073578
experiencia2	11.25	0.088928
publico_expe	4.65	0.215277
invmls2	4.35	0.229820
educacion	4.12	0.242464
publico	3.38	0.296212
jefe_hogar	2.17	0.460075
nor_sierra	1.43	0.697613
ctro_sierra	1.36	0.732632
mujer	1.34	0.745353
sur_sierra	1.31	0.764812
guayaquil	1.25	0.798448
amazonia	1.17	0.856774
indigena	1.06	0.942268
migracion	1.03	0.974282
<b>Mean VIF</b>	<b>3.56</b>	

Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

El Factor Inflación de la Varianza (VIF) de las variables experiencia y experiencia al cuadrado es mayor a 10, por tanto, existe problema de multicolinealidad<sup>69</sup> en el modelo; no obstante, las variables que evidencian este problema son significativas y el regresor experiencia al cuadrado se construye a partir de la variable experiencia; por ende, no es necesario corregir la multicolinealidad. Además que la teoría económica, en especial las ecuaciones de ingresos propuestas por Mincer (1974) requieren de la presencia de dicha variable; pues una de las razones es que se verifica la concavidad negativa de la variable experiencia al cuadrado ya que como se mencionó los ingresos de los individuos crecen hasta un punto determinado de la experiencia y luego empieza su decrecimiento.

<sup>69</sup> Para mayor información del tema véase Castro (2010), pp. 71-92.

### Prueba de heterocedasticidad

Según Castro (2010), los primeros detectores de heterocedasticidad son los gráficos de dispersión y los gráficos de residuos; por consiguiente, anteriormente ya se analizó el gráfico de residuos en función de las predicciones y se concluyó que no existen indicios de heterocedasticidad; es decir, al parecer la varianza es constante; sin embargo, para confirmar esta afirmación se realiza la prueba de Breusch-Pagan.

La hipótesis nula de la prueba de Breusch-Pagan / Cook-Weisberg hace referencia a que el modelo es homocedástico; mientras que la hipótesis alternativa se refiere a la heterocedasticidad del modelo; analíticamente<sup>70</sup> se tiene:

$$H_0: \sigma_i^2 = \sigma^2$$

$$H_1: \sigma_i^2 = \sigma^2 h(\alpha_2 z_{i2} + \dots + \alpha_n z_{in})$$

Los resultados del modelo son:

#### Prueba de Heterocedasticidad de Breusch-Pagan / Cook-Weisberg

Ho: Varianza constante

chi2(1)	0.52
Prob> chi2	0.4726

Según esta prueba, se confirma que el modelo multiplicativo planteado no presenta problemas de heterocedasticidad, ya que no se rechaza la hipótesis nula de homogeneidad de los errores.

Notas:

- ✓ No se realiza la validación del modelo aditivo, el motivo es que el modelo multiplicativo considera interacciones entre las variables explicativas que provienen del modelo aditivo; por consiguiente, se analizará las estimaciones del modelo multiplicativo por ser más completo.

<sup>70</sup> Para ampliar información sobre la prueba de Breusch-Pagan, véase Breusch, T. S., and A. R. Pagan. 1979. A simple test for heteroscedasticity and random coefficient variation, The Journal of Econometric Society, Econometrica 47, pp 1287–1294.

- ✓ Según Galindo (2006), la suposición de que los residuos estén normalmente distribuidos no es necesaria para la estimación de los parámetros de regresión pero si es indispensable para la construcción de las pruebas de hipótesis e intervalos de confianza de los parámetros. No obstante, la prueba F es robusta para la no normalidad y si consideramos que la prueba ji cuadrado es una aproximación de la prueba F; entonces, en el presente trabajo de investigación la normalidad no sería un problema. Además, otra de las razones que justifica que la normalidad no es un problema es que el número de observaciones utilizadas es suficientemente grande (5542 individuos); ya que por el teorema del límite central, según Galindo (2006)<sup>71</sup>, se dice que cuando n tiende al infinito, la distribución de la variable tiende hacia una distribución normal estándar.

## **ANEXO- B Validación de los modelos representantes de los sectores público y privado**

El presente trabajo de investigación realiza estimaciones de las ecuaciones de ingresos por separado para cada sector laboral (público y privado) ya que como se mencionó en el apartado 3.3.3 se requiere analizar la descomposición de ingresos de cada grupo; por tanto, es necesaria la validación de ambas ecuaciones.

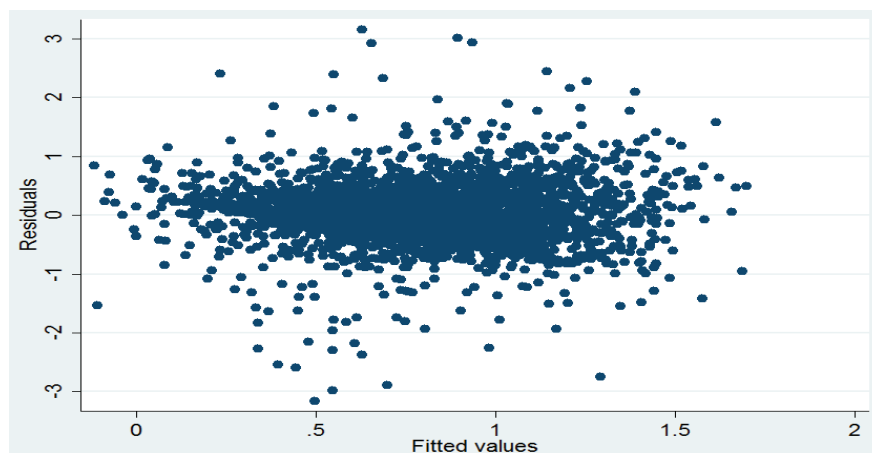
### **Validación del modelo para el sector privado**

A continuación se validarán los supuestos de linealidad, homocedasticidad, normalidad y no correlación de los residuos de la ecuación de ingresos del sector privado:

---

<sup>71</sup> Para obtener mayor información del teorema del límite central véase Galindo, Edwin (2006), “Estadística métodos y aplicaciones para administración e ingeniería”, ProCiencia Editores, Quito, pp. 149-153.

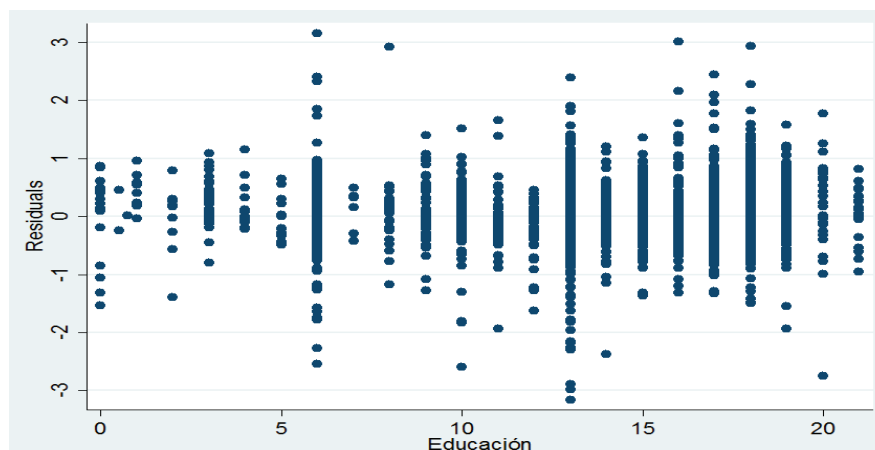
### Gráfico de residuos en función del pronóstico



Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

El gráfico de residuos en función del pronóstico no muestra indicios de heterocedasticidad, ni de falta de linealidad y normalidad. Esto es evidente pues la nube de puntos se encuentra en una franja horizontal simétrica y las observaciones al parecer se encuentran distribuidas al azar. Sin embargo, se observa la existencia de pocos puntos atípicos.

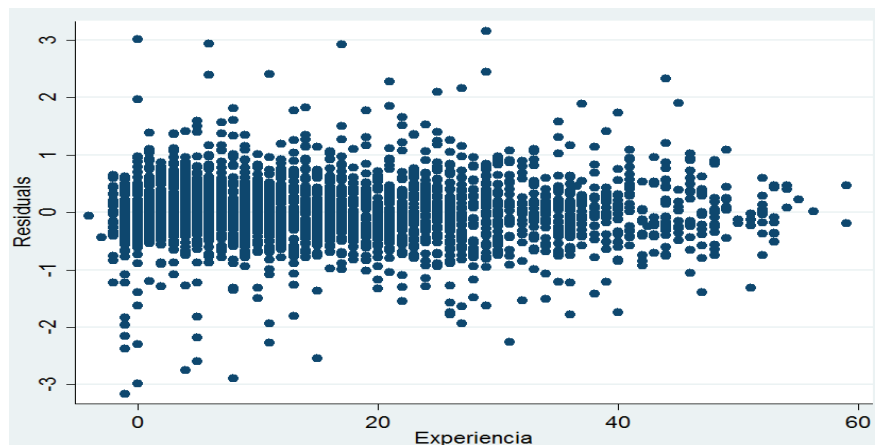
### Gráfico de residuos en función de los años de educación



Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

Aparentemente, el presente gráfico no muestra problemas con el cumplimiento de las hipótesis; pero si hay la presencia de puntos atípicos que se encuentra alrededor del rango -3 y 3.

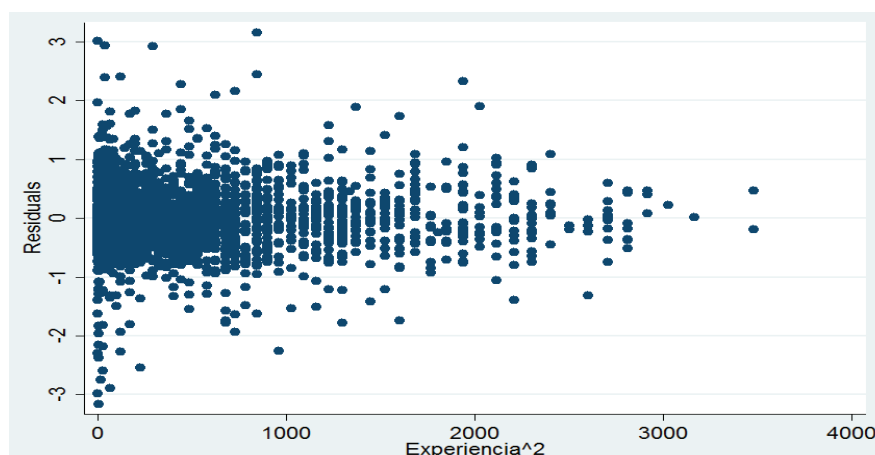
### Gráfico de residuos en función de los años de experiencia



Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

Las observaciones de los residuos en relación con la variable experiencia tienden a distribuirse aleatoriamente en un intervalo de -3 y 3 respecto al eje de las y's, lo que confirmaría que no existe violación de las hipótesis; aunque se puede identificar 4 puntos atípicos.

### Gráfico de residuos en función de los años de experiencia al cuadrado

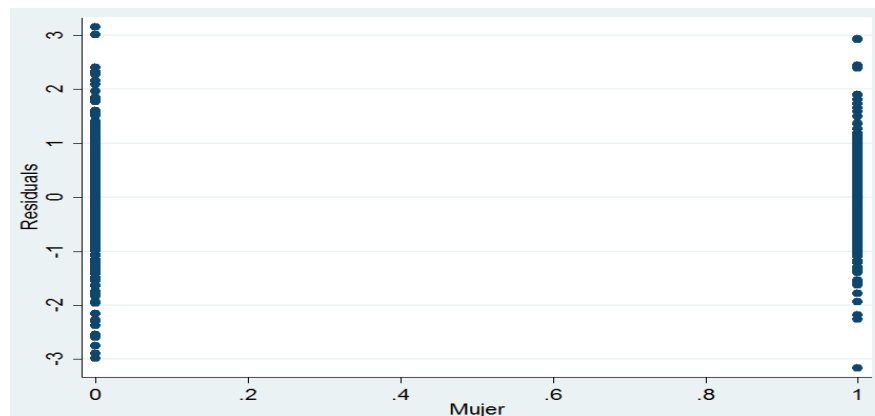


Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

Al parecer este gráfico muestra problemas con la varianza de los residuos, pues se observa problemas de heterocedasticidad y adicionalmente se identificarán ciertos puntos atípicos.



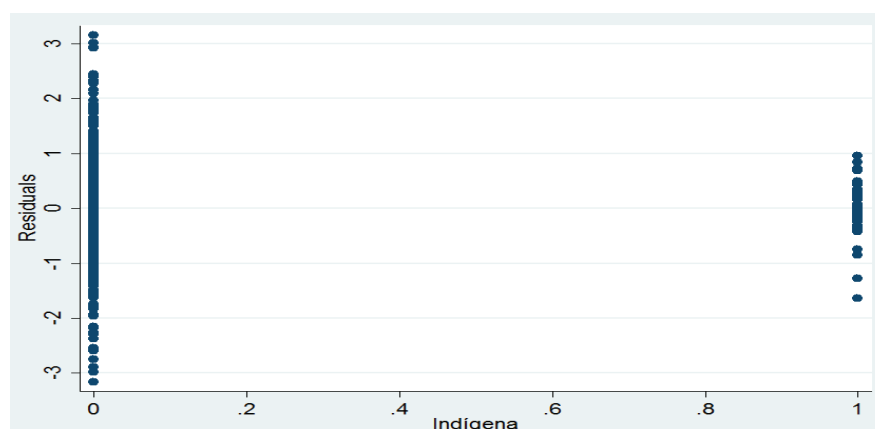
### Gráfico de residuos en función de género



Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

Este gráfico presenta dos poblaciones ya que la variable mujer es dicotómica (que toma el valor de 1 si los individuos son mujeres o 0 si son hombres); por ende, se forman dos columnas verticales de puntos y se puede decir que tanto en la distribución de las observaciones de las mujeres como en la de los hombres existen puntos atípicos pero no hay evidencia de violación de las hipótesis.

### Gráfico de residuos en función de auto identificación étnica



Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

Similarmente al gráfico anterior se forman dos columnas verticales porque la variable indígena es dicotómica (toma el valor de 1 para la población que se auto identifica como indígena y cero caso contrario) y se identifican 4 puntos atípicos solamente en

la población no indígena. Además, al parecer en este gráfico las hipótesis se cumplen.

### Gráfico de residuos en función de relación de parentesco



Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

Las dos poblaciones que se observan en el gráfico representan a los individuos que se consideran jefes de hogar (1) y a los que no se consideran jefes de hogar (0). En ambos grupos se observan puntos atípicos; no obstante, el incumplimiento de las hipótesis no se evidencia.

### Gráfico de residuos en función de la situación migratoria

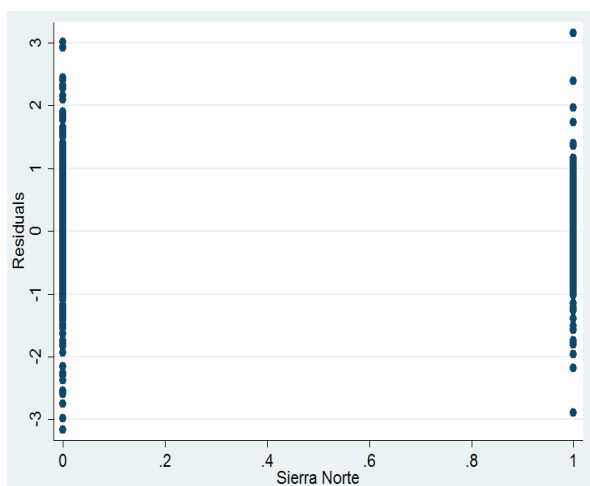


Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

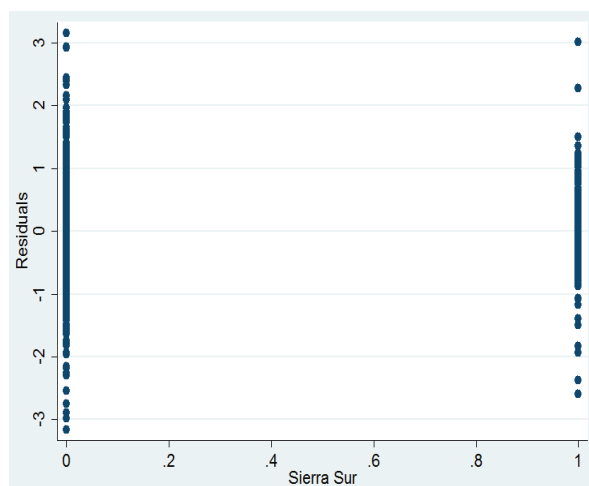
Este gráfico al igual que los anteriores presenta dos poblaciones, una de los individuos que han migrado de su ciudad de origen y otra conformada por las

personas que no han migrado. Superficialmente, no se han incumplido con las hipótesis pero si existen puntos atípicos en el grupo de personas que no han migrado.

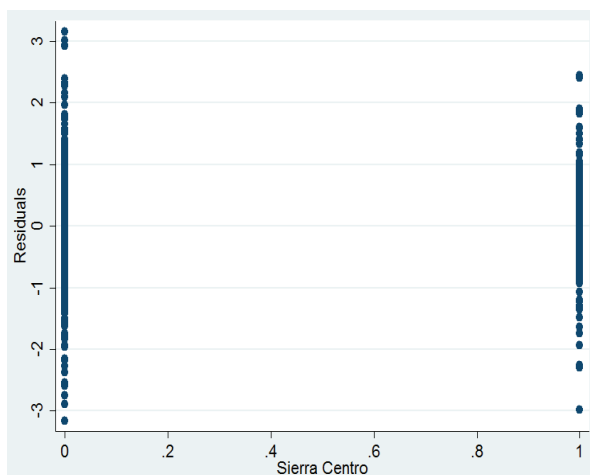
**Gráfico de residuos en función de la ubicación sierra norte**



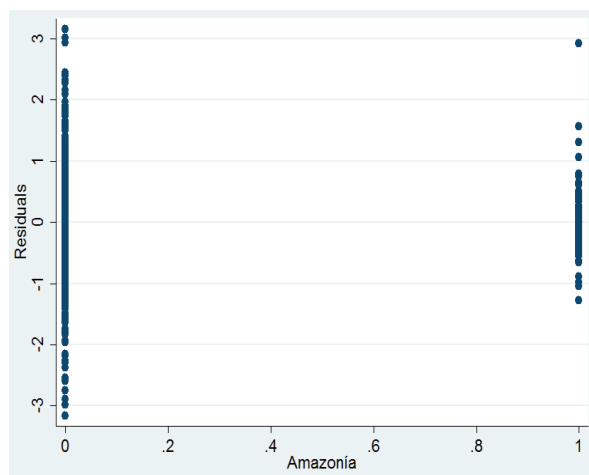
**Gráfico de residuos en función de la ubicación sierra sur**



**Gráfico de residuos en función de la ubicación sierra centro**



**Gráfico de residuos en función de la región amazónica**

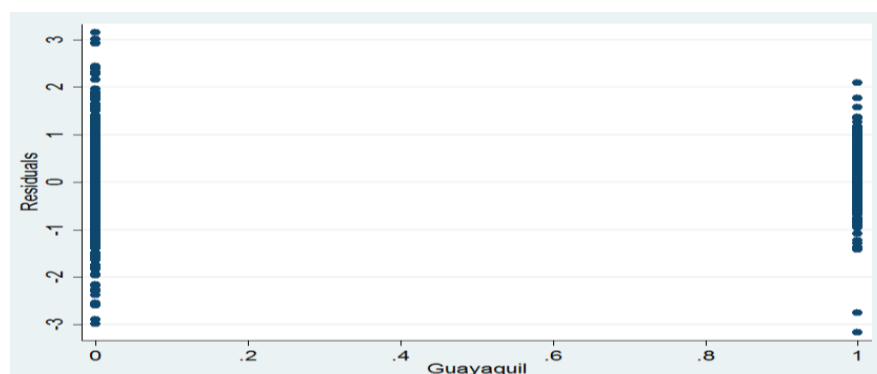


Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

Las variables Sierra Norte, Sierra Centro, Sierra Sur y Amazonía representan la variable de localización geográfica y se observa en sus 4 gráficos que al parecer las

hipótesis de linealidad, no correlación, homocedasticidad y normalidad se cumplen aunque existan ciertos puntos atípicos.

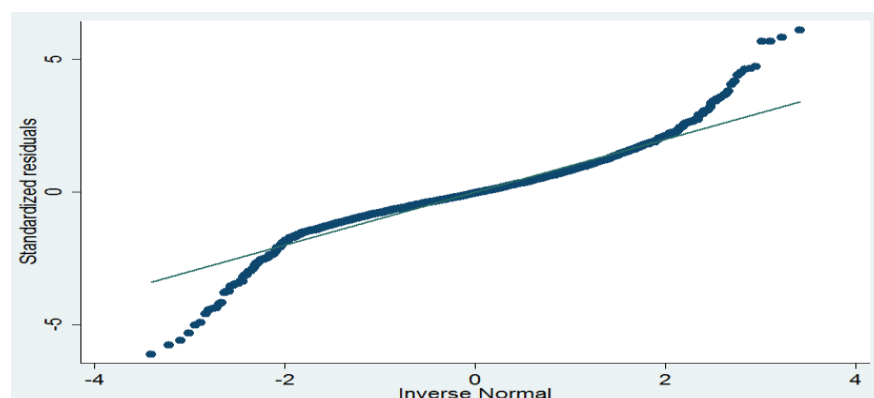
### Gráfico de residuos en función de la ubicación ciudad de Guayaquil



Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

En este gráfico se identifican puntos atípicos tanto para los individuos que viven en la ciudad de Guayaquil como para los que habitan en el resto del país; por consiguiente, dichos puntos serán estudiados más adelante.

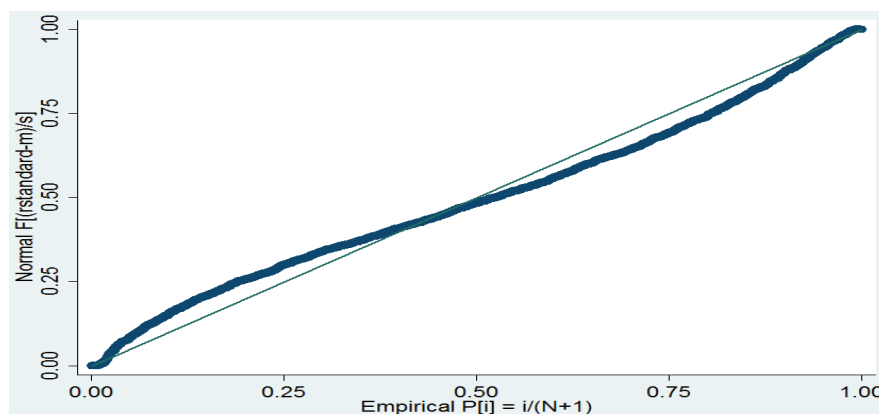
### Gráfico de cuantiles de los residuos estandarizados frente a los cuantiles de la distribución normal



Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

El gráfico de cuantiles no evidencia que los residuos sigan una ley normal; pues presenta en los extremos derecho e izquierdo observaciones que se alejan de la recta de probabilidad normal.

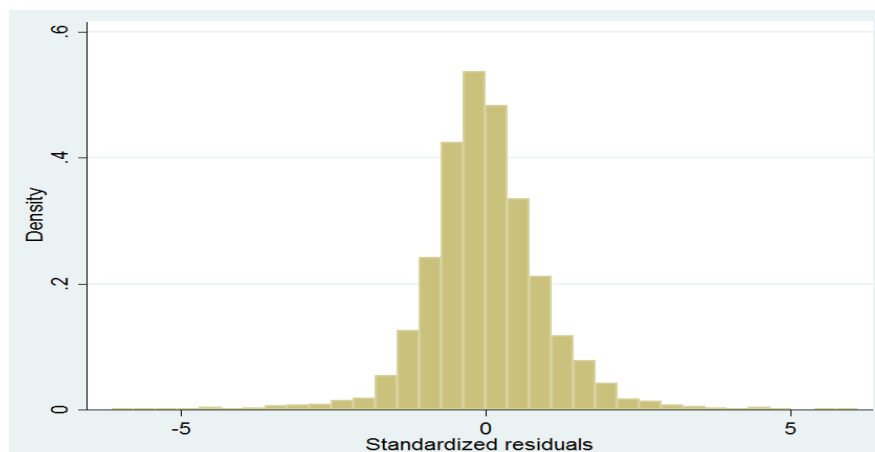
### Gráfico de percentiles de los residuos estandarizados frente a los percentiles de la distribución normal



Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

Este gráfico al igual que el anterior evidencia que los residuos no siguen una ley normal ya que existen puntos que no se ajustan a la recta de la función de distribución normal.

### Histograma de residuos estandarizados



Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

El histograma expone una curtosis alta y aparentemente existe asimetría por lo que se sospecha que los residuos siguen una ley diferente a la normal. Por lo tanto, a continuación se realiza la pruebas respectivas de asimetría y curtosis.

### Prueba de apuntamiento y asimetría de los residuos estandarizados

Variable	Pr(Asimetría)	Pr(Curtosis)	Prob>chi2
rstandard	0.0461	0.0000	0.0000

Fuente: ENEMDU, diciembre 2011

Elaboración: La autora

En base a la interpretación que Carrillo (2009) expone, se acepta la hipótesis nula de existencia de simetría ( $H_0: CA=0$ ) y se rechaza la hipótesis de que el coeficiente de apuntamiento ( $CA_p$ ) sea igual a 3; por ende, si solo se acepta una hipótesis y se rechaza la otra, existe la sospecha de que los errores siguen una ley diferente a la normal; es decir se rechaza la hipótesis nula de normalidad de los residuos<sup>72</sup>.

En conclusión, según la interpretación que Carrillo (2008) plantea, en el histograma y en el gráfico de probabilidad normal se observa que un 95% de los residuos estandarizados se encuentran entre -2 y 2, y el 99% entre -3 y 3; sin embargo, la prueba de hipótesis nula de normalidad de los residuos se rechaza; entonces, es necesario identificar datos atípicos ya que en el gráfico de probabilidad normal existen observaciones al inicio y al final del gráfico que no se ajustan a la recta. Por consiguiente, la eliminación de los puntos problema podrían corregir la normalidad de los residuos.

Para la identificación de los puntos atípico, palanca e influyentes, el paquete estadístico Stata (2012) propone ciertas reglas y formas de cálculo ya mencionadas en la validación del modelo multiplicativo. Seguidamente se presenta un resumen de los puntos mencionados.

---

<sup>72</sup> Por las razones ya mencionadas en la validación del modelo multiplicativo; no es necesario que se cumpla con la hipótesis de normalidad de los residuos; sin embargo, se trata de corregir la normalidad eliminando ciertos puntos problema.

### Puntos palanca

Variable	Obs	Mean	Std. Dev.	Min	Max
ppalanca	761	.0082434	.0048775	.0048779	.0289459

Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

### Puntos influyentes

Variable	Obs	Mean	Std. Dev.	Min	Max
pinfluye	295	.0023336	.0027858	.0007245	.0309801

Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

### Puntos atípicos

Variable	Obs	Mean	Std. Dev.	Min	Max
patipico	25	4.093706	.9511671	3.057595	6.160289

Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

### Puntos palanca influyentes no atípicos

Variable	Obs	Mean	Std. Dev.	Min	Max
ppalanca	153	.0100138	.0064398	.0048956	.0289459

Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

### Puntos palanca influyentes atípicos

Variable	Obs	Mean	Std. Dev.	Min	Max
ppalanca	9	.0076936	.0033385	.0050614	.0132011

Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

Con el objetivo de que se cumpla la hipótesis de normalidad de los residuos, se eliminarán de la muestra las observaciones que se muestran a continuación:

### Puntos palanca influyentes y atípicos

Observación	$h_{ii}$	$D_j$	$r_j$
706.	.0050912	.0071001	4420631
1197.	.0050614	.0048787	3671618
1198.	.0059709	.0055134	3591635
2491.	.0057372	.0038428	3057595
3469.	.0051443	.0125578	5862582
4253.	.0132011	.0309801	5723376
5376.	.0093093	.0091949	3708938
5487.	.0130107	.0087809	3057955
5778.	.0067163	.0104944	467721

Fuente: ENEMDU, diciembre 2011

Elaboración: La autora

Con el nuevo conjunto de datos, el modelo estimado por el método de Heckman (1979) es así:

Source	SS	df	MS	Number of obs = 3072		
Model	254.734286	13	19.5949451	F( 13, 3058) = 77.35		
Residual	774.633264	3058	.25331369	Prob > F = 0.0000		
Total	1029.36755	3071	.335189694	R-squared = 0.2475		
				Adj R-squared = 0.2443		
				Root MSE = .5033		

w	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
educacion	.0618353	.0041332	14.96	0.000	.0537312	.0699394
experiencia	.0168638	.0022982	7.34	0.000	.0123577	.0213699
experiencia2	-.0001244	.0000494	-2.52	0.012	-.0002214	-.0000275
mujer	-.1076093	.020755	-5.18	0.000	-.1483044	-.0669142
indigena	-.101126	.0666423	-1.52	0.129	-.2317943	.0295423
jefe_hogar	.0730667	.0305973	2.39	0.017	.0130733	.1330601
migracion	.0376888	.0258708	1.46	0.145	-.013037	.0884147
nor_sierra	.0645752	.0251111	2.57	0.010	.015339	.1138115
ctro_sierra	.035837	.0299642	1.20	0.232	-.022915	.094589
sur_sierra	.0657581	.0297744	2.21	0.027	.0073782	.124138
amazonia	.1855695	.0546244	3.40	0.001	.0784653	.2926737
guayaquil	.1296396	.0301141	4.30	0.000	.0705936	.1886856
invmills5	-.2622118	.130582	-2.01	0.045	-.5182492	-.0061745
_cons	-.2049193	.0999027	-2.05	0.040	-.4008025	-.0090361

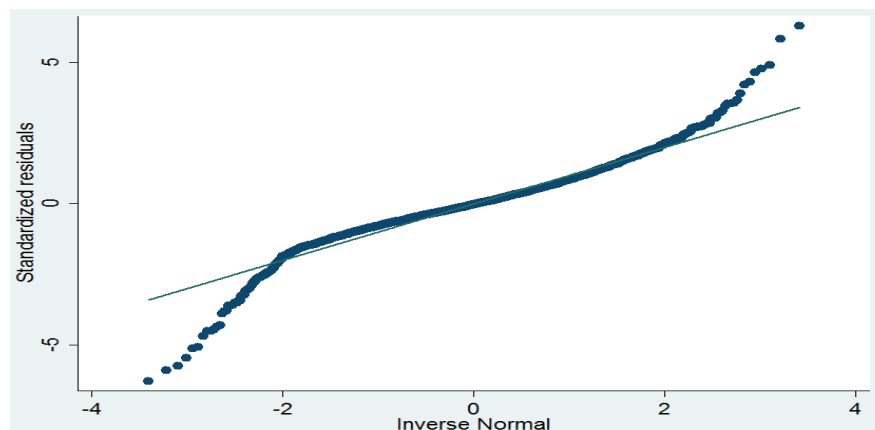
Fuente: ENEMDU, diciembre 2011

Elaboración: La autora



A continuación, se valida las hipótesis del nuevo modelo así:

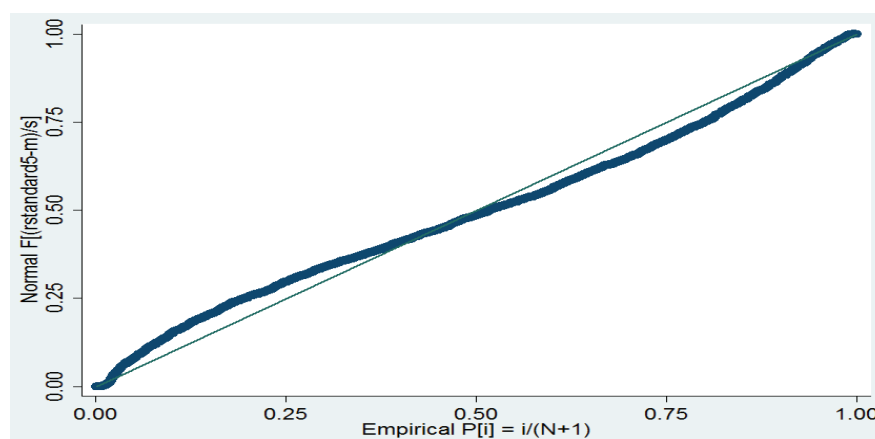
### Gráfico de cuantiles de los residuos frente a los cuantiles de la distribución normal



Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

No se observa una mejora con respecto al gráfico de cuantiles anterior, pues los puntos a los extremos se mantienen alejados de la recta de la función normal. Entonces al no haber un cambio, el modelo anterior representaría al sector privado.

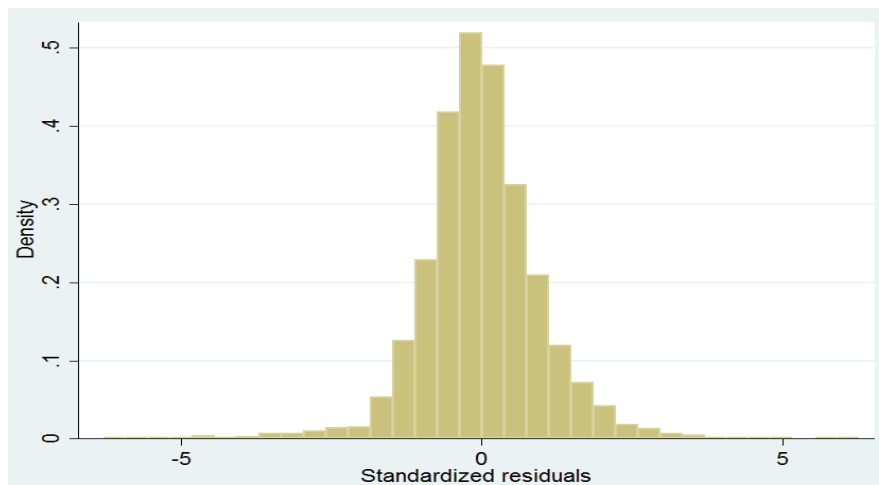
### Gráfico de percentiles de los residuos frente a los percentiles de la distribución normal



Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

Lo mismo ocurre con el presente gráfico, no se observa ningún cambio favorable con respecto al gráfico del modelo anterior. Los puntos siguen alejados de la recta normal.

### Histograma de residuos estandarizados



Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

La asimetría y la curtosis siguen siendo un problema en el histograma; en consecuencia, para verificar esta afirmación se validan estas hipótesis como sigue:

#### Prueba de apuntamiento y asimetría de los residuos estandarizados

Variable	Pr(Asimetría)	Pr(Curtosis)	Prob>chi2
rstandard	0.0031	0.0000	0.0000

Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

Estas pruebas indican que se rechazan las hipótesis nulas de curtosis y simetría; consecuentemente, se rechaza la hipótesis nula de que los residuos sigan una ley normal. En conclusión, el primer modelo es el que mejor se ajusta a una ley normal.

Luego de la validación de la hipótesis de normalidad se confirman las hipótesis de no correlación y homocedasticidad del modelo.

#### Prueba de multicolinealidad

La fórmula estadística para el cálculo del Factor inflación de la Varianza (VIF) se expone en la validación del modelo multiplicativo y para el modelo del sector privado este estadístico es:

Variable	VIF	1/VIF
experiencia	10.84	0.092235
experiencia2	9.63	0.103855
educacion	4.07	0.245471
invmills4	4.07	0.245796
jefe_hogar	2.70	0.370535
nor_sierra	1.43	0.700500
guayaquil	1.31	0.763596
ctro_sierra	1.29	0.775981
sur_sierra	1.29	0.777883
mujer	1.27	0.790304
amazonia	1.09	0.920429
migracion	1.07	0.935030
indigena	1.03	0.969895
Mean VIF	3.16	

Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

El VIF indica que no existen problemas de multicolinealidad en el modelo puesto que todos sus valores son menores a 10 aunque el VIF de la variable experiencia es 10.84; sin embargo, no es un VIF alto.

### Prueba de heterocedasticidad

Para comprobar la heterocedasticidad del modelo, Breusch-Pagan / Cook-Weisberg plantean una hipótesis nula de que el modelo es homocedástico y una hipótesis alternativa que hace referencia a la heterocedasticidad del modelo; analíticamente se tiene que:

$$H_0: \sigma_i^2 = \sigma^2$$

$$H_1: \sigma_i^2 = \sigma^2 h(\alpha_2 z_{i2} + \dots + \alpha_n z_{in})$$

Para el modelo que representa al sector privado los resultados son:

#### Prueba de Heterocedasticidad de Breusch-Pagan / Cook-Weisberg

Ho: Varianza constante

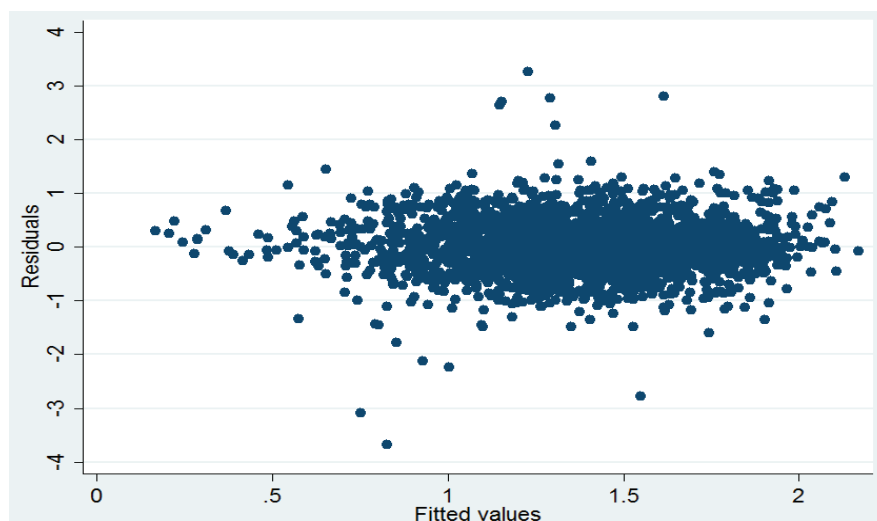
chi2(1)	30.3
Prob > chi2	0.0000

Si bien el gráfico de residuos en función del pronóstico no muestra violación de las hipótesis, estos resultados implican que se rechaza la hipótesis nula evidenciando problemas de heterocedasticidad. Este problema puede ser causado por variables omitidas en el modelo ya que dado las limitaciones de la encuesta utilizada no se han considerado variables fundamentales como las que representan el historial familiar.

### Validación de modelo para el sector público

Se debe verificar que se cumplan los supuestos de los residuos para la ecuación de ingresos que representa al sector público; por tanto, primeramente se estudian los gráficos de los residuos así:

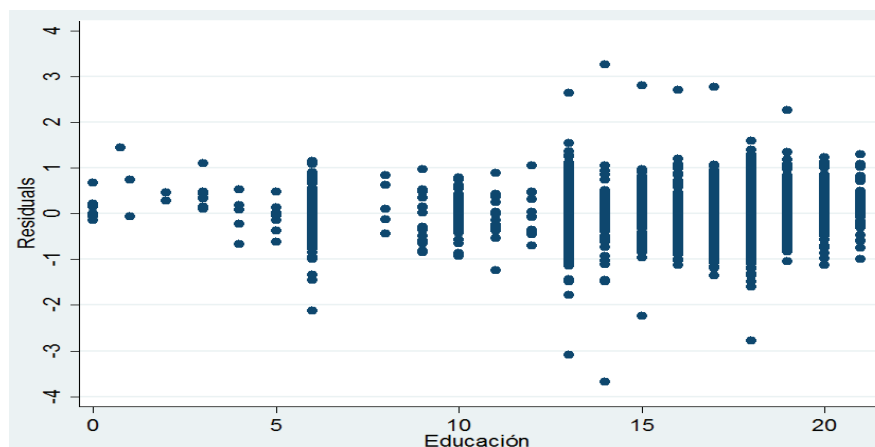
**Gráfico de residuos en función del pronóstico**



Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

El gráfico presenta una nube de puntos cuyas observaciones se encuentran distribuidas al azar en una franja horizontal simétrica entre -3 y 3; por ende, no hay evidencia de que la varianza no sea constante y de que no se cumplan los supuestos de linealidad, normalidad y no correlación de los residuos, aunque se identifican 3 puntos atípicos.

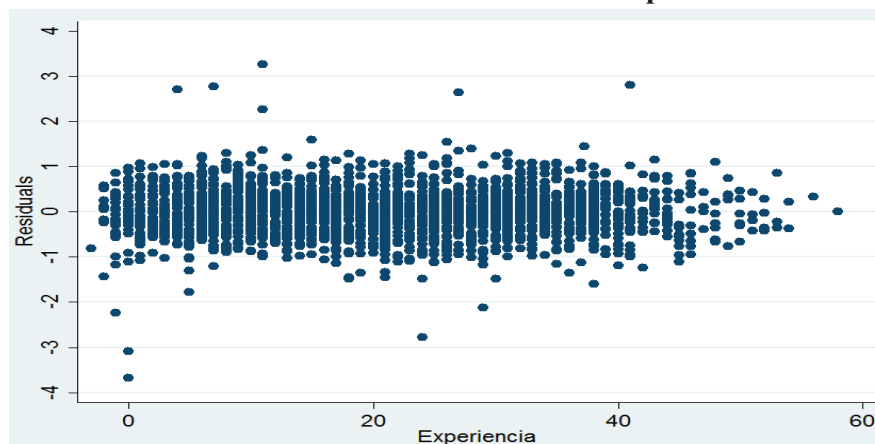
**Gráfico de residuos en función de la educación**



Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

El gráfico de los residuos con respecto a la educación no presenta problemas relacionados con la violación de las hipótesis mencionadas en el gráfico anterior; no obstante, también se observa la presencia de 3 puntos atípicos.

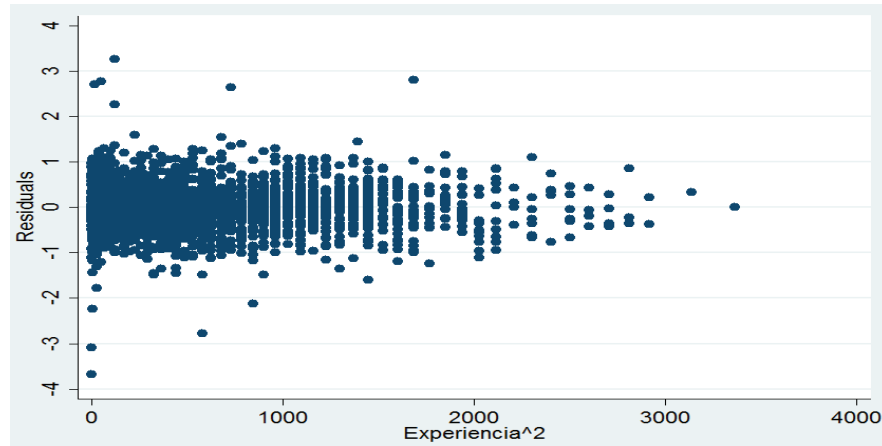
**Gráfico de residuos en función de la experiencia**



Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

Este gráfico indica que todas las hipótesis planteadas se cumplen a pesar de la existencia de pocos puntos atípicos; pues la nube de puntos cumple con el requisito de que el 99% de las observaciones se encuentren distribuidas aleatoriamente en el intervalo  $[-3, 3]$ .

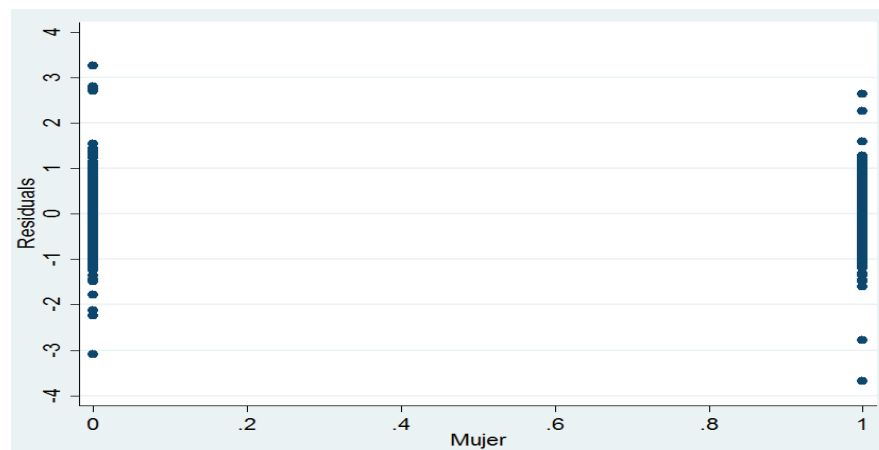
### Gráfico de residuos en función de la experiencia al cuadrado



Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

El gráfico muestra problemas con la varianza ya que al parecer esta no es constante; además, se encuentran 3 puntos atípicos que podrían originar la no normalidad de los errores.

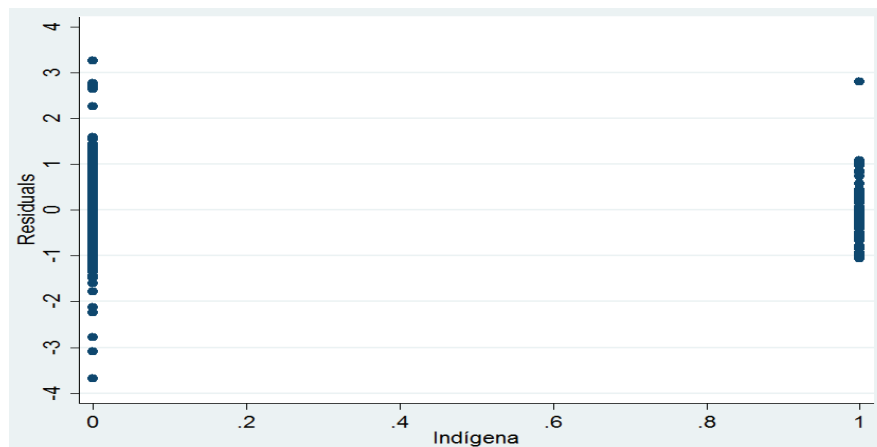
### Gráfico de residuos en función de género



Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

Se identifican dos poblaciones, una para los hombres (0) y otra para las mujeres (1). Aparentemente, no se observan problemas con las hipótesis pero si hay presencia de puntos atípicos en ambos grupos.

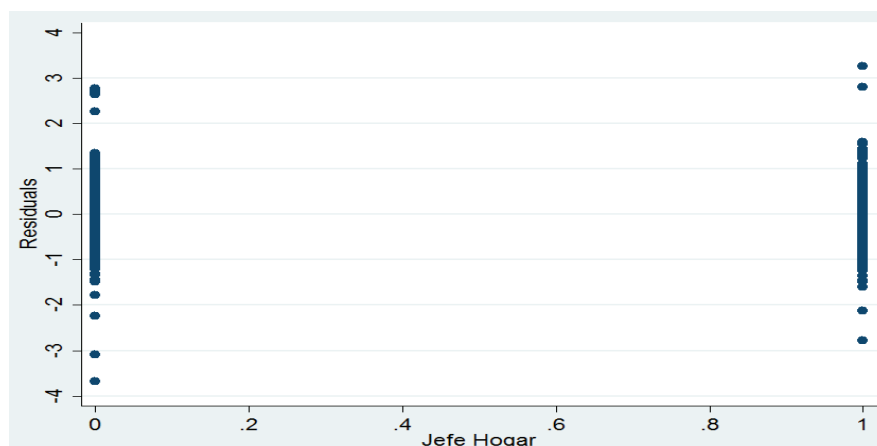
### Gráfico de residuos en función de auto identificación étnica



Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

Los puntos atípicos en este gráfico corresponde solamente a los individuos auto identificados como no indígenas; sin embargo, se puede decir que no existen problemas con el incumplimiento de las hipótesis.

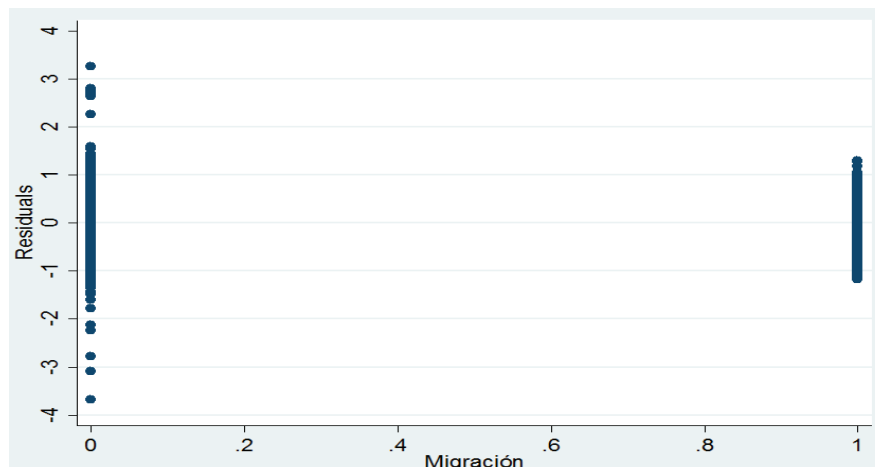
### Gráfico de residuos en función de relación de parentesco



Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

No hay indicios de que las hipótesis no se cumplen a pesar de que existen puntos atípicos en el grupo de los individuos que declararon ser jefes de hogar y también en el grupo de individuos que declararon no ser jefes de hogar.

### Gráfico de residuos en función de la situación migratoria

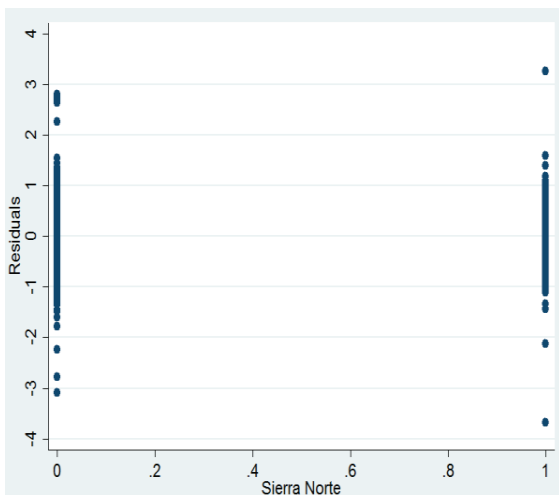


Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

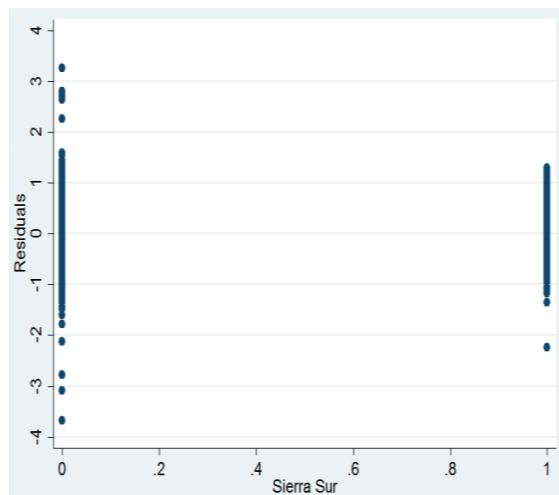
Al parecer de acuerdo al gráfico las hipótesis de los residuos se cumplen y aunque existen ciertos puntos atípicos pueden estos no ser un problema; por consiguiente, más adelante deberán ser estudiadas dichas observaciones para eliminarlas o para que se mantengan en el conjunto de datos.



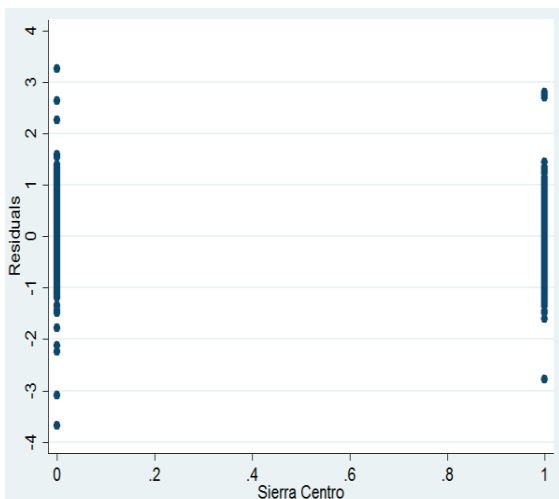
**Gráfico de residuos en función de la ubicación sierra norte**



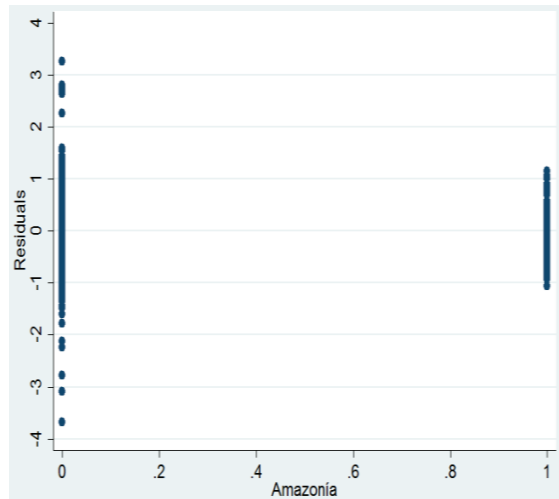
**Gráfico de residuos en función de la ubicación sierra sur**



**Gráfico de residuos en función de la ubicación sierra centro**



**Gráfico de residuos en función de la región amazónica**

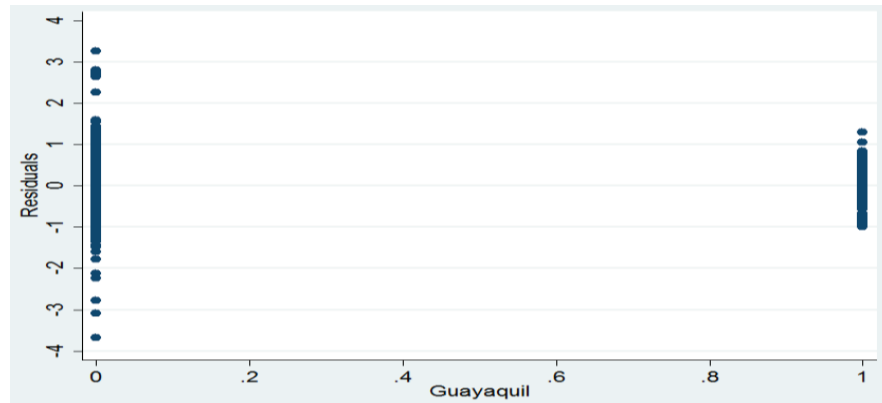


Fuente: ENEMDU, diciembre 2011

Elaboración: La autora

Las variables independientes de estos gráficos representan a la variable localización geográfica. Los gráficos no muestran indicios de violación de las hipótesis; solamente exhiben ciertos puntos atípicos que de ser necesario serán eliminados.

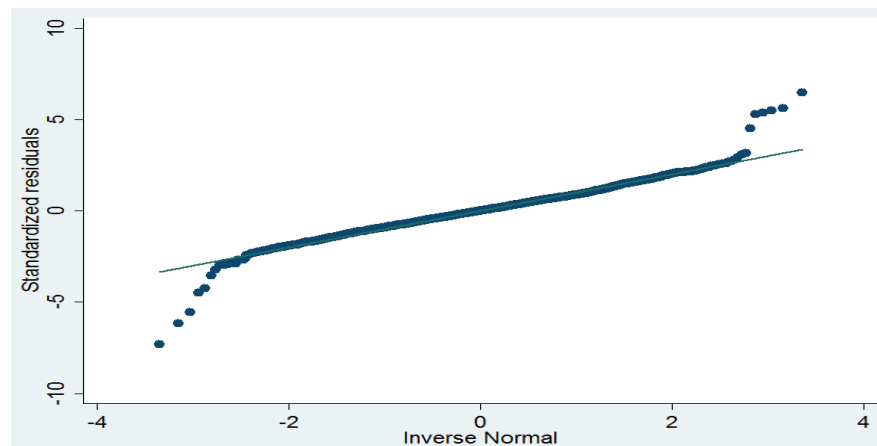
### Gráfico de residuos en función de la ubicación ciudad de Guayaquil



Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

La población de individuos que no habitan en la ciudad de Guayaquil está conformada por 3 puntos atípicos que podrían influenciar en la normalidad de los errores. Por otra parte el gráfico presentado no muestra incumplimiento del resto de hipótesis de los residuos.

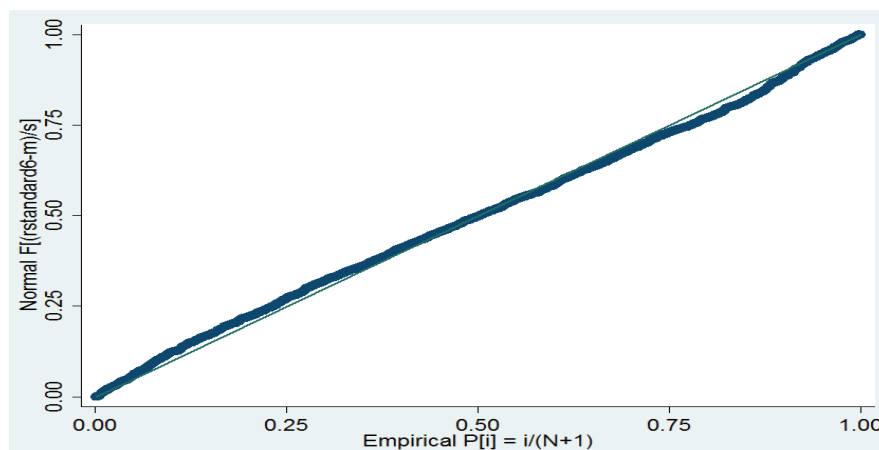
### Gráfico de cuantiles de los residuos estandarizados frente a los cuantiles de la distribución normal



Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

El gráfico presenta puntos desviados de la recta normal por lo que se sospecharía que los residuos no siguen una ley normal.

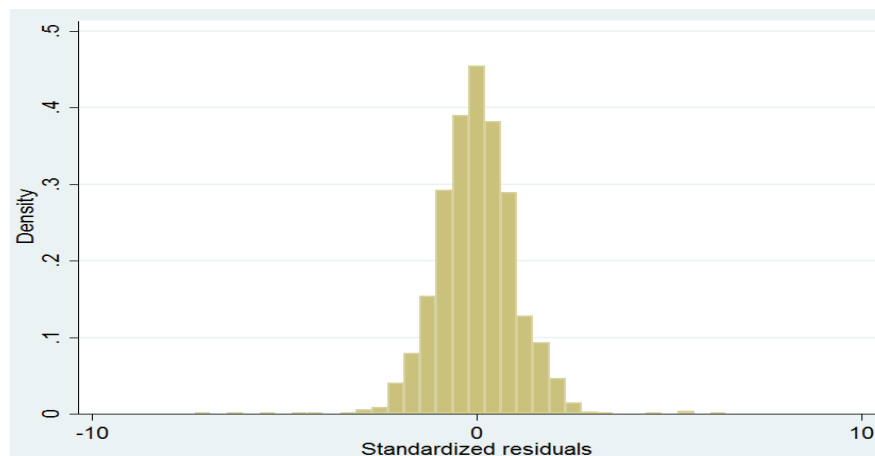
### Gráfico de percentiles de los residuos estandarizados frente a los percentiles de la distribución normal



Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

La mayoría de puntos se ajustan a la recta de la función de distribución normal; no obstante lo deseable sería que todos los puntos estén sobre dicha recta.

### Histograma de residuos estandarizados



Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

En el histograma se observa la existencia de simetría; pero hay problemas con la curtosis ya que el apuntamiento al parecer es mayor a 3; por consiguiente, se realizarán las pruebas de hipótesis de normalidad correspondientes.

### Prueba de apuntamiento y asimetría de los residuos estandarizados

Variable	Pr(Asimetría)	Pr(Curtosis)	Prob>chi2
rstandard	0.6501	0.0000	0.0000

Fuente: ENEMDU, diciembre 2011

Elaboración: La autora

En base a los resultados de las pruebas de hipótesis de normalidad de los residuos, en general, se rechaza la hipótesis nula de que los residuos sigan una ley normal<sup>73</sup>; sin embargo, las pruebas de simetría y curtosis indican que se debe rechazar la hipótesis nula de que el coeficiente de apuntamiento sea igual a 3 y se debe aceptar la hipótesis nula de existencia de simetría.

El problema que se presenta con la normalidad de los residuos podría ser causado por la existencia de puntos atípicos, palanca o influyentes; entonces es necesario analizar dichas observaciones y para ello se utiliza los criterios y fórmulas del programa estadístico Stata (2012) ya descritos en la validación del modelo multiplicativo.

Seguidamente se expone un resumen de los puntos palanca, influyentes y atípicos:

#### Puntos palanca

Variable	Obs	Mean	Std. Dev.	Min	Max
ppalanca	1054	.0085486	.0039095	.0048791	.0325257

Fuente: ENEMDU, diciembre 2011

Elaboración: La autora

#### Puntos influyentes

Variable	Obs	Mean	Std. Dev.	Min	Max
pinfluye	336	.0019702	.0027441	.0007226	.0356082

Fuente: ENEMDU, diciembre 2011

Elaboración: La autora

<sup>73</sup> Como ya se mencionó en la validación del modelo multiplicativo; no es necesario que se cumpla con la hipótesis de normalidad de los residuos; sin embargo, se intentará corregir la normalidad eliminando los puntos palanca, influyente y atípicos.

### Puntos atípicos

Variable	Obs	Mean	Std. Dev.	Min	Max
patipico	8	4.897783	1.223885	3.080539	6.540533

Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

### Puntos palanca influyentes no atípicos

Variable	Obs	Mean	Std. Dev.	Min	Max
ppalanca	217	.0093593	.0041747	.0048877	.0267054

Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

### Puntos palanca influyentes atípicos

Variable	Obs	Mean	Std. Dev.	Min	Max
ppalanca	2	.0103093	.0074738	.0050245	.015594

Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

Luego de identificar los puntos que podrían ser un problema para la normalidad de los errores, se eliminan de la muestra las observaciones que se muestran a continuación:

### Puntos palanca influyentes y atípicos

Observación	$h_{ii}$	$D_j$	$r_j$
484.	.015594	.0356082	5.645066
3310.	.0050245	.0036284	3.177488

Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

Con el nuevo conjunto de datos se obtiene las siguientes estimaciones por el método de Heckman (1979):

Source	SS	df	MS			
Model	225.028139	13	17.3098569	Number of obs =	2459	
Residual	607.756123	2445	.248571012	F( 13, 2445) =	69.64	
				Prob > F =	0.0000	
				R-squared =	0.2702	
				Adj R-squared =	0.2663	
Total	832.784263	2458	.33880564	Root MSE =	.49857	

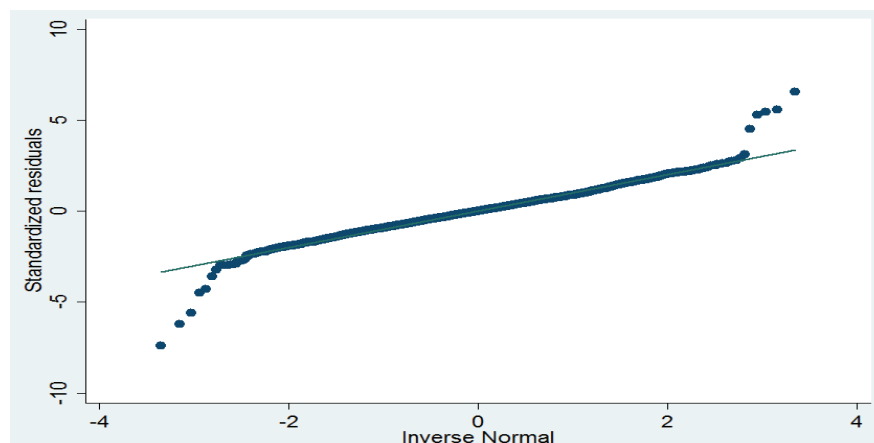
  

w	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
educacion	.0614708	.0049519	12.41	0.000	.0517604	.0711811
experiencia	.0168695	.0031582	5.34	0.000	.0106765	.0230624
experiencia2	-.0000633	.0000649	-0.98	0.329	-.0001906	.000064
mujer	-.1477133	.0240763	-6.14	0.000	-.1949254	-.1005012
indigena	-.1568327	.0557882	-2.81	0.005	-.2662297	-.0474356
jefe_hogar	.0149133	.0279503	0.53	0.594	-.0398955	.069722
migracion	.109666	.0303695	3.61	0.000	.0501135	.1692186
nor_sierra	.0817167	.030587	2.67	0.008	.0217376	.1416958
ctro_sierra	.0823932	.0279181	2.95	0.003	.0276476	.1371388
sur_sierra	.121932	.0313098	3.89	0.000	.0605355	.1833285
amazonia	.1664045	.0438661	3.79	0.000	.080386	.252423
guayaquil	.1068151	.0496621	2.15	0.032	.009431	.2041991
invmills8	-.1755004	.0727491	-2.41	0.016	-.3181567	-.0328442
_cons	.128236	.1185046	1.08	0.279	-.1041439	.3606158

Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

Es necesario validar el nuevo modelo así:

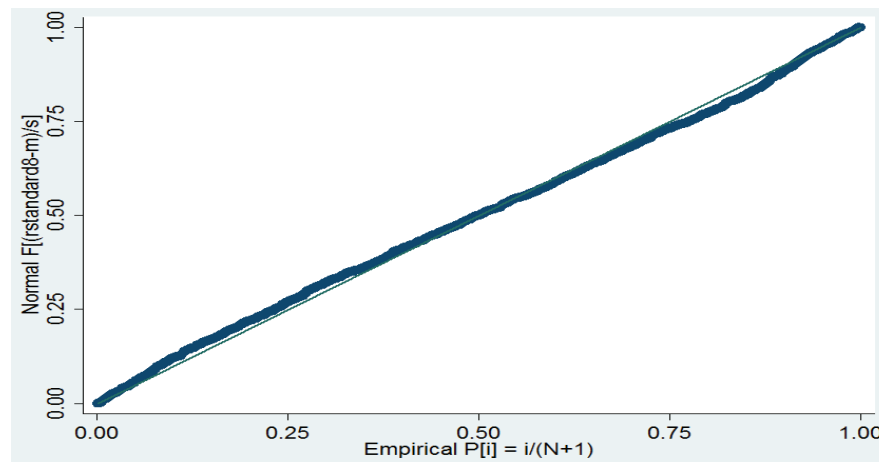
### Gráfico de cuantiles de los residuos estandarizados frente a los cuantiles de la distribución normal



Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

No se observa un cambio con respecto al gráfico de cuantiles anterior; pues se mantienen ciertos puntos alejados de la recta de distribución normal.

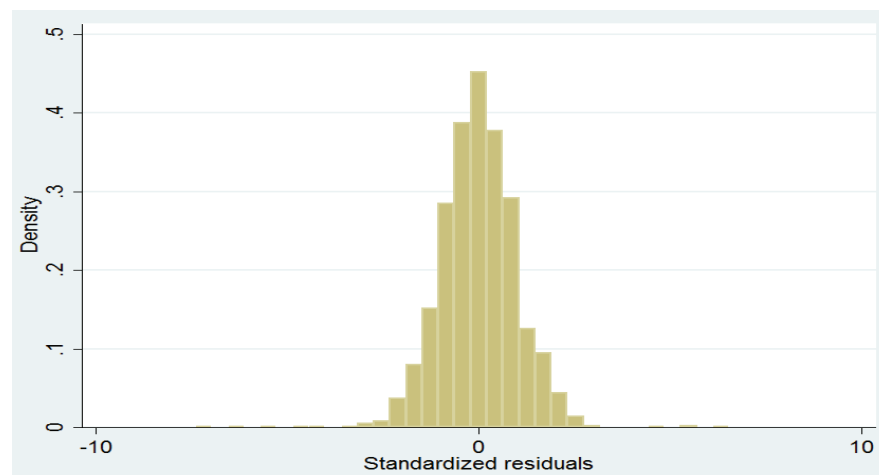
### Gráfico de percentiles de los residuos estandarizados frente a los percentiles de la distribución normal



Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

Tampoco en el gráfico de percentiles existe la percepción de un cambio significativo ya que los puntos en su totalidad no se ajustan a la recta de referencia.

### Histograma de residuos estandarizados



Fuente: ENEMDU, diciembre 2011  
Elaboración: La autora

El actual histograma parece no haber cambiado ya que el problema de apuntamiento observado permanece y la simetría aparentemente no ha mejorado. Paralelamente, se presenta las pruebas de hipótesis de normalidad de los residuos:

### Prueba de apuntamiento y asimetría de los residuos estandarizados

Variable	Pr(Asimetría)	Pr(Curtosis)	Prob>chi2
rstandard	0.2813	0.0000	0.0000

Fuente: ENEMDU, diciembre 2011

Elaboración: La autora

De forma similar al anterior modelo se rechaza la hipótesis nula de normalidad de los residuos; igualmente se rechaza la hipótesis nula de que el coeficiente de apuntamiento sea igual a 3 y se acepta la hipótesis de que el coeficiente de asimetría sea igual a cero. En conclusión, el primer modelo será el que mejor represente al sector público ya que su probabilidad de asimetría (0.6501) es mayor a la probabilidad de asimetría del segundo modelo (0.2813).

Luego de probar la normalidad de los residuos se debe comprobar que no existan problemas de multicolinealidad y de heterocedasticidad.

### Prueba de multicolinealidad

Para el cálculo del factor inflación de la varianza (VIF) se utiliza el programa estadístico Stata (2012) y su fórmula se expresó en la validación del modelo multiplicativo. Por lo tanto, los VIF del modelo que representa al sector público son los siguientes:

Variable	VIF	1/VIF
experiencia	15.67	0.063828
experiencia2	13.80	0.072476
invmills7	3.76	0.266203
educacion	3.36	0.298059
jefe_hogar	1.92	0.520304
mujer	1.43	0.698369
ctro_sierra	1.43	0.701728
nor_sierra	1.42	0.704470
sur_sierra	1.34	0.746479
amazonia	1.25	0.799359
guayaquil	1.13	0.882940
indigena	1.12	0.893789
migracion	1.03	0.970914
Mean VIF	3.74	

Fuente: ENEMDU, diciembre 2011

Elaboración: La autora



Todos los VIF son menores a 10 a excepción de los VIF de las variables experiencia y experiencia al cuadrado; sin embargo al tener dichas variables unos VIF no muy altos se puede concluir que el modelo no tiene problemas de multicolinealidad.

### **Prueba de heterocedasticidad**

Para la prueba de heterocedasticidad del modelo, existen dos hipótesis planteadas por Breusch-Pagan / Cook-Weisberg que son: una hipótesis nula de que el modelo es homocedástico y una hipótesis alternativa que indica la heterocedasticidad del modelo; analíticamente se tiene que:

$$H_0: \sigma_i^2 = \sigma^2$$

$$H_1: \sigma_i^2 = \sigma^2 h(\alpha_2 z_{i2} + \dots + \alpha_n z_{in})$$

Para el modelo que representa al sector público los resultados son:

#### Prueba de Heterocedasticidad de Breusch-Pagan / Cook-Weisberg

Ho: Varianza constante

chi2(1)	25.14
Prob > chi2	0.0000

Los resultados indican que se debe rechazar la hipótesis nula; es decir existen problemas de heterocedasticidad; sin embargo, el gráfico de residuos en función del pronóstico no muestra evidencia de incumplimiento con las hipótesis. No obstante, como se mencionó en la validación del modelo que representa al sector privado, este problema puede ser causado por variables omitidas en el modelo ya que dado las limitaciones de la encuesta no se han considerado variables fundamentales como las que representan el historial familiar.