

29
UNIVERSIDAD AUTONOMA DE NUEVO LEON

FACULTAD DE ECONOMIA



TESIS

**EN OPCION AL TITULO DE
LICENCIADO EN ECONOMIA**

**Análisis de Descomposición por Factores Componentes de
la Desigualdad del Ingreso: Un Estudio para el Mercado Laboral
del Area Metropolitana de Monterrey**

LUIS CUAUHEMOC GONZALEZ RIVERA

MONTERREY, N. L.

AGOSTO DE 1999

TL

HD6996

Z8

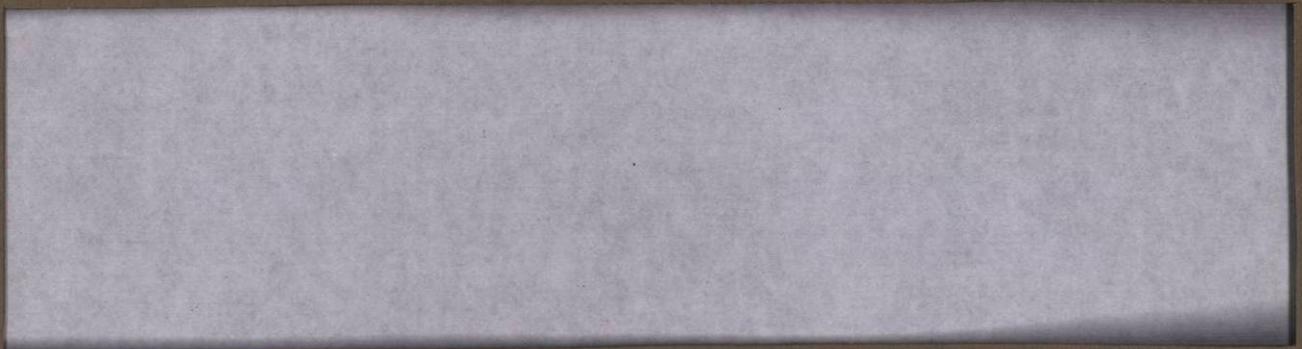
N84

1999

c.1



1080092546



**Análisis de Descomposición por Factores Componentes de
la Desigualdad del Ingreso: Un Estudio para el Mercado Laboral
del Area Metropolitana de Monterrey**

LUIS CUAUHEMOC GONZALEZ RIVERA

MONTERREY, N. L.

AGOSTO DE 1999

TL
HD6996
.Z8
N84
1999



Dedicado a la mujer más maravillosa que he conocido, a la mujer cuyo ejemplo, amor y comprensión me formó e hizo de mi todo lo que soy, y pueda llegar a ser, la mujer que sin vacilar dedicó su tiempo, trabajo e inteligencia en pro de su familia. A la mujer que nos irradió alegría y optimismo a todos los que la conocimos y nos enseñó a valorar lo bello de la vida, dedicado a la memoria de mi hermosa mamá, Arcelia Rivera, gracias siempre.

Agradezco de todo corazón el apoyo total que recibí por parte de mi familia a lo largo de los años de carrera, los cuales culminan con la presentación de este trabajo. Muchas gracias a mi papá el Capitán Luis González Galindo, a mis hermanas Xochitl y Citlali, a mi querida abuelita "Toñita," a mi tía Rosario y a todos y cada uno de mis tíos, primos y amigos de la familia. Gracias por su consejo y amistad sincera.

Agradecimientos

Gracias a todos mis profesores, quienes inculcaron en mi el deseo de superación y excelencia; cada uno de ellos me ayudó a crecer en el ámbito académico y a madurar como persona. Me gustaría agradecer en especial a mi asesor de Tesis el Dr. Jorge N. Valero Gil, quien me animó en la realización del proyecto y creyó en mí, así como al Dr. Pedro A. Villezca Becerra y al Dr. José Alfredo Tijerina Guajardo, por sus comentarios sobre el trabajo y por formar parte del Comité de Evaluación en mi Examen Profesional; al Lic. Jorge Meléndez Barrón, al Dr. Marco Vinicio Gómez Meza y al Dr. Dennis Young de la Universidad Estatal de Arizona por su apoyo y sugerencias en la realización del presente trabajo y a lo largo de todos los años de carrera, gracias por su ejemplo.

Sería injusto no agradecer a mis compañeros de generación su apoyo, tanto en las buenas como en las malas, por lo que les doy las gracias por ser magníficos amigos y amigas, por su originalidad y talento nato, gracias a todos.

Índice

	<i>Página</i>
Capítulo I. Introducción	1
Capítulo II. Marco Teórico	8
2.1 La descomposición de la desigualdad del ingreso	
2.1.1 <i>Propiedades de las diferentes medidas de desigualdad</i>	9
2.1.2 <i>Propiedades de la descomposición</i>	11
2.2 Metodología	13
Capítulo III. Variación del Ingreso en el Área Metropolitana de Monterrey, Un Modelo Lineal	
3.1 Los datos	16
3.2 Exploración y análisis de la base de datos	17
3.3 Cuantificación de los efectos personales	22
3.4 Análisis de varianza, modelo de efectos principales	
3.4.1 <i>Introducción al análisis de varianza</i>	24
3.4.2 <i>Aplicación del análisis de varianza al ingreso individual</i>	26
3.4.3 <i>Análisis de varianza dentro de grupos educativos</i>	32
Capítulo IV. Variación del Ingreso en el Área Metropolitana de Monterrey, Un Modelo con Interacciones	
4.1 Interpretación de las interacciones en el modelo	36
4.2 Análisis de varianza, modelo completo con interacciones	38
Capítulo V. Conclusiones y Análisis Final	51
Bibliografía	54
Apéndice 1. Gráficas adicionales	
Apéndice 2. Resultados de pruebas estadísticas	
Apéndice 3. Análisis de Varianza para México, por regiones y nivel educativo	

Capítulo I. Introducción

Los juicios sobre la importancia de varias influencias sobre la desigualdad del ingreso tienen una larga historia. A lo largo del tiempo se ha vuelto cada vez más común relacionar estos juicios a índices sumarios de la desigualdad, e intentar descomponer el valor de la desigualdad agregada dentro de las contribuciones de los componentes relevantes. En los últimos años se ha seguido un método que observa la desigualdad dentro de un país y/o ciudades importantes. Este tipo de análisis usa una variedad de metodologías y descompone la desigualdad por regiones económicas (por ejemplo, zona noroeste Vs. sureste), entre estados o entre ciudades, por sector económico (Manufacturas Vs. Servicios), fuentes del ingreso (ingreso laboral Vs. capital Vs. tierra Vs. transferencias) o características familiares (incluyendo atributos de los trabajadores, sus trabajos y otras consideraciones de localización). Este modo de inferencia es potencialmente valioso para entender la estructura de la desigualdad y la desigualdad a través del tiempo y para identificar los factores más importantes que causan que algunos individuos o familias reciban ingresos más altos que otros.¹

Este trabajo propone que se deben verificar empíricamente no grandes esquemas interpretativos, sino hipótesis más modestas sobre dos áreas específicas: empleo y distribución del ingreso. Urge hacerlo antes de que estas ideas caigan en el olvido.

La definición y medición de la desigualdad constituye un área de investigación de gran porvenir para los economistas latinoamericanos, quienes requieren medidas de la desigualdad del ingreso que reflejen las características de las variables económicas que dan lugar a distintas distribuciones del ingreso. Los coeficientes de desigualdad más populares, los de Gini y de Pareto, no cumplen con este requerimiento. El uso de medidas como los coeficientes de Atkinson, Theil y el Análisis de Varianza (ANOVA, por sus siglas en inglés) constituyen un avance en esta dirección. Dicho de otra forma, es necesario ligar significativamente las variables que, según la teoría económica, son

¹ Véase Shorrocks (1982) y (1983).

responsables de la distribución del ingreso, con la teoría estadística de medición de la desigualdad.²

Se discute si los economistas deben considerar tópicos sobre la distribución. Varias personas piensan que no. Las nociones concernientes a la distribución “correcta” del ingreso son juicios de valor y no hay una forma “científica” para resolver diferencias en asuntos de ética. Por lo tanto, algunos piensan que la discusión sobre problemas de distribución está en detrimento de la objetividad en economía y los economistas deberían restringirse a sí mismos a analizar sólo los aspectos de eficiencia de los asuntos y problemas sociales.

Este punto de vista tiene un problema, la Teoría Económica del Bienestar indica que la eficiencia por sí misma no puede ser utilizada para evaluar una situación dada. Un criterio aparte del de eficiencia debe ser usado para comparar asignaciones alternativas de recursos. Claro que uno puede decir que solamente la eficiencia importa, pero esto en sí mismo es un juicio de valor.

Aunque el principal objetivo del presente estudio es la desigualdad económica en la distribución del ingreso, es importante tener en mente que ésta constituye sólo una pequeña parte del gran problema de la desigualdad en un país en vías de desarrollo como México. De paralela importancia o incluso mayor están las desigualdades de poder, prestigio, estatus, reconocimientos, condiciones de trabajo, grado de participación y muchas otras dimensiones del problema del desarrollo relacionadas con la autoestima y la libertad de elección. Pero como en muchas relaciones sociales, uno no puede realmente separar las manifestaciones económicas de las no-económicas de la desigualdad. Cada una refuerza a la otra en un complejo y común proceso interrelacionado de causa y efecto.

Los objetivos de la presente Tesis son medir la importancia relativa de los efectos personales y del tipo de actividad económica, en que una persona se desenvuelve, sobre la variación del ingreso entre los habitantes pertenecientes a la población económicamente activa del Área Metropolitana de Monterrey (AMM). Y determinar los distintos factores que explican las divergencias en ingresos dentro de segmentos relativamente homogéneos

² Véase Shorrocks (1982) y Fields (1980a).

de la población. Se le da un énfasis especial a la búsqueda de diferencias en ingreso entre hombres y mujeres a través de distintas clasificaciones, tales como los grupos de edad, el tipo de empleo, etc.

El estudio se enfoca en la aplicación de un método que nos ayude a entender el *por qué* de los diferenciales en los ingresos, entre las personas del mercado laboral de una de las zonas urbanas más importantes del país: Monterrey. Se desea conjuntar una serie de variables para poder medir su importancia relativa en la determinación de los ingresos de la Población Económicamente Activa (PEA) del Área Metropolitana y, a su vez, estudiar las interrelaciones existentes entre éstas mismas. Adicionalmente, se investiga la variabilidad del ingreso dentro de grupos homogéneos de la población, segmentándola por grados educativos.

El análisis nos permite investigar las posibles interacciones existentes entre cada par de factores estudiados. Esto es, podemos investigar la influencia que ejerce un factor explicativo en las diferencias internas en ingresos que presente un segundo factor. Las interacciones se denominan de “dos vías,” es decir, por pares de factores explicativos. Están diseñadas para probar si se mantiene, o no, la estructura de los diferenciales de los ingresos entre los diversos niveles de un factor, observados “dentro” de un nivel (o subdivisión) de un segundo factor explicativo, cuando se examinan todos los niveles del segundo factor. De este modo, se probaría estadísticamente si los diferenciales, o no-diferenciales, entre diversas categorías clasificativas se mantienen en todo momento.

Se decidió seguir esta línea de investigación por la riqueza en información proporcionada sobre el mercado laboral del ÁMM, aprovechando una metodología muy utilizada en estudios para países desarrollados, pero con un escaso número de publicaciones para el caso latinoamericano y países en vías de desarrollo en general. La literatura sobre el tópico se enfoca en el estudio de índices que sintetizan la desigualdad en un número y analiza su comportamiento a través del tiempo. O bien, se estudian las propiedades específicas de variables explicativas comúnmente empleadas en los estudios de la desigualdad del ingreso, por ejemplo, algunos estudios se concentran en el cálculo de las tasas de rendimiento de la escolaridad, utilizando diferentes criterios para definir la propia escolaridad, entre éstos se hallan Rupert, Severance y Turner (1996) y Kenny

(1983); para el caso del AMM se pueden consultar: Meléndez (1995, 1997a y 1997b), Camero (1995), Garro y Llamas (1995), López (1983) y Valdez (1995). Los estudios abarcan diversos aspectos relacionados con la educación y sus tasas de rendimiento, buscando diferencias por género, carreras universitarias, grados alcanzados o la interrelación existente entre educación y capacitación en el trabajo. Los modelos empleados utilizan, principalmente, el análisis de regresión, con algunas variantes dependiendo las necesidades del estudio.

En la presente Tesis se propone una metodología de investigación muy utilizada en las ciencias sociales, pero, poco aprovechada en la investigación económica, el Análisis de Varianza. A través de este tipo de análisis podemos conocer la importancia relativa de todos los factores estudiados en el modelo, sin necesidad de enfocarnos en un sólo factor. Una de las mayores ventajas de las técnicas del análisis de varianza es que indica la importancia cuantitativa de cada variable explicativa. Luego entonces, podemos aprender de las descomposiciones por el análisis de varianza cuánta diferencia (en porcentajes por categoría explicativa) se produce, si una persona labora en un sector económico en lugar de otro o si tiene un año adicional de estudios.

Es conveniente, en este punto, adelantar algunos de los resultados más interesantes de la investigación. Los factores a estudiar son: sexo, grupos de edad, último grado de estudio alcanzado, antigüedad en el último empleo, situación sindical, tipo de empleo, estado civil y actividad económica. Una vez estudiado y contrastado el “poder explicativo” de los ocho factores considerados en los modelos de análisis de varianza, resulta ser la *educación* la variable con mayor proporción de la variabilidad explicada de los ingresos entre las personas de la PEA del AMM, en todos los casos. La categoría educación posee, por mucho, el mayor peso explicativo, mientras que le siguen en forma descendente el tipo de empleo, la edad, la actividad económica en donde se labora, la antigüedad en el último empleo y el estado civil. Casi una quinta parte de los diferenciales en ingresos puede ser explicada únicamente por el grado de estudios alcanzado por una persona. Este es un punto interesante, dado que una de las variables a la que se le podría atribuir un poder explicativo importante (la actividad económica donde labora una persona) tiene un menor peso en relación con el factor grado de estudio alcanzado, a

pesar de que en la actividad económica se están estudiando y apreciando variaciones entre 12 sectores.³

Las variables sexo y situación sindical no poseen un efecto estadísticamente significativo, es decir, no se observan diferencias significativas en los ingresos promedio entre hombres y mujeres. Por tanto, parece no existir discriminación por género para la PEA del AMM, en cuanto a ingresos por hora se refiere.

A su vez, no se detectaron interacciones significativas entre el factor educativo y los factores actividad económica, estado civil, antigüedad en el último empleo, el sexo ni con los grupos de edad, por tanto, se concluye que la importancia del capital humano a lo largo de todas las actividades económicas, los estados civiles, los rangos de antigüedad y de edad y entre hombres y mujeres es reconocida y premiada de manera muy similar dentro de las ramas clasificatorias de cada uno de estos factores. De igual forma, se encontraron interacciones significativas entre la educación y los factores tipo de empleo y situación sindical, esperado en el primer caso por estar muy ligados la gama de tipos de empleos disponibles con el grado educativo, pero, no así en el segundo, dado que no se habían detectado diferencias entre los ingresos promedio de las personas sindicadas de las no sindicadas. Esperábamos que se mantuviera esta relación de no-diferencias dentro del factor situación sindical a través de los grados educativos, como de hecho sucedió entre las clasificaciones internas de los demás factores (demostrándose por medio de las interacciones correspondientes entre el factor situación sindical y cada uno de los otros factores), sin embargo, se encontró que sólo se mantenía para los grupos con educación superior y para aquellos con educación primaria, beneficiándose de la membresía del sindicato solamente las personas con educación secundaria y media superior. Este es un resultado conjunto del modelo de análisis de varianza completo con interacciones y un análisis de varianza para cada uno de los cuatro grupos educacionales.

Algo similar se observa en el caso del factor actividad económica, el cual no interactúa con los grupos de edad, el estado civil, la situación sindical ni con el sexo, indicando que la estructura de ingresos de estos factores (diferencias en los primeros dos

³ Existen varios trabajos donde se remarca la importancia del sector económico o el giro de la empresa en la determinación de los ingresos individuales, incluso se les atribuye un poder explicativo mayor que a variables de capital humano. Véase Groshen (1991) y Gatica, Mizala y Romaguera (1995).

casos, no-diferencias en los últimos) se mantiene constante a través de las ramas de la actividad económica. Mientras que en los casos del tipo de empleo y la antigüedad, estos sí interactúan con la actividad económica, por tanto, al cambiar de rama de la actividad económica cambian las diferencias internas de los factores. Es decir, en algunos sectores económicos se compensa más los años de antigüedad en el empleo, en relación a otros en donde los diferenciales son menores, de manera parecida sucede en el caso del tipo de empleo. Punto de interés para aquellos que buscan empleo en algún área de la economía, basando su elección en los ingresos esperados por el tipo de empleo o el perfil de ingresos según la antigüedad en el empleo.

Finalmente tenemos el factor sexo. A través del análisis de las interacciones de éste con cada uno de los factores restantes concluimos que no se presentan diferencias en los ingresos promedio de los hombres y las mujeres pertenecientes a la PEA del AMM. Comparándose los ingresos promedio por género dentro de cada uno de los otros factores analizados, no se observaron diferencias significativas entre éstos, al no interactuar el sexo con algún otro factor, por tanto, se confirman los resultados previos del análisis de varianza de efectos principales.

De esta manera, el análisis presentado abre una gama de posibilidades de investigación en el área económica, ayudándonos a entender y comprobar lo que la Teoría Económica ha venido estudiando desde años atrás.

La organización del trabajo es la siguiente: en el Capítulo II se analiza la justificación teórica del trabajo, así como los modelos a utilizarse en la parte empírica. En éste se explican las diversas formas en que se pueden descomponer los factores determinantes de la desigualdad y los argumentos técnico-teóricos que nos llevan a escoger al análisis de varianza como el procedimiento a seguir en la descomposición de los diferenciales en ingresos por factores componentes. Este procedimiento se aplicará a los datos muestrales de la Encuesta de Migración, Empleo y Salarios en el Área Metropolitana de Monterrey, Marzo-Abril de 1996.

En el Capítulo III se exploran algunas diferencias en la distribución del ingreso a través de diversas categorías explicativas y se inicia con el análisis empírico utilizando los modelos de regresión y de análisis de varianza. En el Capítulo IV se extiende el modelo

de análisis de varianza presentado en el capítulo precedente, enfatizando los resultados propios del modelo completo, con interacciones. Para concluir en el Capítulo V con algunas inferencias sobre el trabajo empírico desarrollado.

Capítulo II. Marco Teórico

"Mientras que un hombre individual es un rompecabezas (irresoluble) en el agregado él se convierte en una certeza matemática. Uno nunca puede, por ejemplo, predecir qué hará cualquier hombre, pero uno puede decir con precisión lo que un número promedio, hará. Los individuos varían, pero los porcentajes permanecen constantes." Sherlock Holmes en Sir Arthur Conan Doyle "The Sign of Four."

El hecho de que trabajadores similares reciban diferentes salarios ha tradicionalmente atraído la atención de los investigadores en Economía Laboral.

El trabajo empírico de investigación inicial sobre este tópico comienza en los 1950's y 1960's, básicamente para países desarrollados, tomando a la industria como la unidad de análisis. Las investigaciones que surgieron sobre este tema trataron con los efectos de varias características de las industrias, como nivel de beneficios, el grado de concentración, grado de sindicación entre trabajadores, el producto, la tecnología y el tamaño de planta sobre el sueldo promedio.⁴

Sin embargo, gracias a desarrollos teóricos nuevos y al acceso a mejores fuentes de información después se cambió el énfasis de la investigación, en contra de lo ya hecho sobre la desigualdad del ingreso (visto éste como salarios y sueldos). El cual tradicionalmente había sido explicado básicamente por factores de demanda (características de la industria), en los 1970's comenzó a ser medido en términos de características de la oferta labora (modelos de capital humano).

Subsecuentemente, en los 1980's una nueva generación de estudios aparecieron enfatizando características de la industria o una compañía en específico, como factores explicativos de la dispersión en salarios, pero al mismo tiempo controlando los factores del lado de la oferta (características del trabajador).

Estos nuevos estudios utilizan datos más desagregados a nivel de una compañía o del propio trabajador, con dotaciones de capital humano de los trabajadores siendo conocidos como un factor central en la desigualdad del ingreso. En otras palabras la atención está enfocada sobre la hipótesis que trabajadores con dotaciones de capital humano idénticas reciben diferentes remuneraciones, dependiendo dónde trabajen éstos.

⁴ Véase las investigaciones clásicas de Slichter (1950) y Master (1969).

Más recientemente, ha habido un renovado interés en el tópico de la dispersión de los salarios, asociados con estudios por Krueger y Summers (1988), Groshen (1991), entre otros. La mayor contribución de estas investigaciones ha sido relacionar evidencia empírica con desarrollos teóricos nuevos. Los trabajos relacionan la existencia de los diferenciales en salarios a la hipótesis de la eficiencia - salario.

La literatura sobre diferenciales en ingreso ha florecido en años recientes, aunque la mayoría se relaciona con los Estados Unidos de América. En Latino América, la investigación sobre los determinantes del ingreso ha estado enfocada principalmente con la estimación de modelos de capital humano y el análisis de segmentos del mercado laboral, aunque recientemente han sido analizados los diferenciales en ingresos inter-industriales para los casos de Brasil, Chile, Uruguay y Venezuela.⁵

2.1 La descomposición de la desigualdad del ingreso

Existen diversas metodologías para descomponer la desigualdad, de las cuales hay cuatro que son utilizadas comúnmente en la literatura sobre este tópico: la descomposición de Gini, la descomposición de Theil, el Análisis de Varianza y la descomposición del Índice de Atkinson.⁶

Dos temas principales surgen al ponderar las ventajas de los diversos procedimientos para la investigación empírica: las propiedades de la medida de desigualdad en sí misma y la riqueza de la información derivada de la descomposición.

2.1.1 Propiedades de las diferentes medidas de desigualdad

Una forma de elegir cuál medida de desigualdad descomponer es considerar la naturaleza básica de la medida. A este respecto, la descomposición de Gini y el Análisis de Varianza aplicado al logaritmo del ingreso sobresalen. El coeficiente de Gini es fácilmente conceptualizado en términos de las curvas de Lorenz. La varianza tiene una base familiar en el análisis estadístico estándar; además, las distribuciones del logaritmo del ingreso son

⁵ Véase Gatica, Mizala y Romaguera (1995).

aproximadamente normales en forma, así la varianza del logaritmo es conceptualmente aceptable. La dificultad con el Índice de Atkinson es que es derivado de un marco de bienestar con el cual muchos investigadores y estudiantes pueden no estar de acuerdo.⁷ Finalmente, el Índice de Theil, no obstante su amplio uso como una medida de desigualdad, no tiene una interpretación clara.

Otro criterio de selección puede ser la utilidad de la medida de desigualdad al realizar comparaciones de desigualdad. Entre los axiomas deseables para este propósito están:⁸

A1. Axioma de la escala irrelevante. Si una distribución es un múltiplo escalar de otra (por ejemplo, el ingreso de todas las personas de una población es $X\%$ el ingreso de otros en una segunda población), entonces las dos distribuciones tienen el mismo grado de desigualdad. Puesto de otra manera, el grado de desigualdad del ingreso se mide con independencia del nivel de ingreso.

A2. Axioma de simetría. Si dos distribuciones del ingreso son idénticas excepto porque diferentes familias reciben mayor ingreso en ambos casos, entonces las dos distribuciones tienen el mismo grado de desigualdad. Esto sigue del principio de tratar a todos los individuos y familias por igual sin importar la distribución del ingreso.

A3. Axioma de la igualación que preserva el rango. Si una distribución se obtiene de otra por la transferencia de una cantidad positiva de ingreso, por parte de una familia relativamente rica a una relativamente pobre preservando su mismo rango dentro de la distribución, entonces la nueva distribución es más equitativa que la anterior.⁹

⁶ Véase Fields (1980a) para una descripción detallada de los índices de desigualdad.

⁷ Atkinson deriva su medida bajo el supuesto que la utilidad es una función solamente del ingreso propio del individuo; sin embargo, la razón por la que muchos investigadores están interesados en la desigualdad se basa en la creencia que las funciones de utilidad también dependen del ingreso de un individuo en relación al de otros.

⁸ Véase Fields y Fei (1978) para un desarrollo axiomático de los juicios concernientes a las comparaciones de la desigualdad y a Shorrocks (1982), para un análisis detallado de los principios de descomposición por factores componentes, donde el autor concluye que “la única regla de descomposición que cumple con todos los supuestos y teoremas analizados es la descomposición natural de la Varianza.”

⁹ Aunque pocas personas estarían en desacuerdo con este axioma, debe notarse que algunos supuestos no triviales sobre la naturaleza de los juicios de bienestar social son necesarios para garantizar que una distribución “más equitativa” sea considerada siempre como “mejor.”

El coeficiente de Gini, el Índice de Theil y el de Atkinson satisfacen estos axiomas. La varianza no cumple con el Axioma de la escala irrelevante. Sin embargo, la escala irrelevante y los otros axiomas se cumplen utilizando la varianza del logaritmo del ingreso. Por lo tanto, las cuatro medidas de desigualdad consideradas aquí cumplen con los criterios axiomáticos para el análisis de descomposición.

Otra consideración de suma importancia es la sensibilidad de las medidas de ingreso a cambios en el ingreso a lo largo de varios puntos dentro de la distribución. Las personas cuyos juicios de valor los inclinen a darle un mayor peso a la posición económica de los pobres querrán escoger la medida que sea más sensible a la desigualdad asociada a los grupos de ingreso. Existen varios estudios donde se demuestra, entre otras cosas, que la varianza del logaritmo del ingreso es más sensible a la desigualdad asociada a la pobreza, el Índice de Theil es más sensible a la desigualdad asociada con los muy ricos y el coeficiente de Gini es más sensible a la desigualdad de aquellos ubicados a la mitad de la distribución. Para los observadores cuyo principal interés esté con la población de menores recursos, el Análisis de Varianza parece una base más apropiada.

2.1.2 Propiedades de la descomposición

La siguiente es una lista de los resultados deseables para los ejercicios de descomposición:

1. Descomponer la desigualdad completa en componentes dentro y entre factores (como las actividades económicas).
2. Medir la contribución bruta de cada factor explicativo de la desigualdad total.
3. Probar la significancia estadística de esos efectos principales.
4. Medir la contribución marginal de cada factor explicativo.
5. Probar la significancia estadística de los efectos marginales.
6. Medir los efectos de las interacciones entre pares de factores explicativos (y combinaciones de alto orden si se necesitan).
7. Probar la significancia estadística de los efectos de las interacciones.

8. Estimar la magnitud de los efectos sobre el ingreso de cada categoría de cada variable explicativa.

El Análisis de Varianza cumple con todas las propiedades; la descomposición de Theil lo hace sólo con la 1, 2, 3, 4 y 6; mientras que las descomposiciones de Gini y Atkinson sólo cumplen con la 1 y 2. En el cuadro 2.1 se presenta un resumen de las propiedades de las medidas de la desigualdad relevantes para la descomposición.

Cuadro 2.1

Medidas de desigualdad

<i>PROPIEDAD</i>	<i>GINI</i>	<i>THEIL</i>	<i>LOG. DE LA VARIANZA</i>	<i>ATKINSON</i>
Descomposición	Sí	Sí	Sí	Sí
Inferencia intuitiva como una medida de la desigualdad	Sí	No	Sí	Dudoso
Justificación axiomática como una medida de desigualdad	Sí	Sí	Sí	Sí
Sensibilidad con respecto a la desigualdad asociada con la pobreza	No	No	Sí	Sí
Clasificación de las contribuciones de los determinantes de la desigualdad	Sí	Sí	Sí	Sí
Estimaciones cuantitativas de la magnitud de los determinantes de la desigualdad	No	No	Sí	No
Pruebas de la significancia estadística de los efectos estimados de varios factores	No	No	Sí	No

Fuente: Gary S. Fields, *"Poverty, Inequality, and Development,"* Cambridge University Press, 1980.

Como se puede observar el logaritmo de la varianza (utilizado en el análisis de varianza, ANOVA, por sus siglas en inglés) es la medida que cumple con todas las propiedades deseables. Sin embargo, esto no significa que sea la medida de desigualdad más empleada por los economistas, como hemos remarcado, de hecho el Índice de Theil es la herramienta de descomposición más utilizada, principalmente por su exactitud, pero, dado que en el presente estudio se examinarán las posibles interacciones entre los factores considerados, seguir el análisis de varianza resulta más conveniente que la metodología del Índice de Theil, por la falta de pruebas estadísticas para los elementos de interacción de esta última. Adicionalmente, el análisis de varianza puede ser complementado por su

paralelo, el análisis de regresión múltiple, permitiendo una mejor comprensión del proceso de determinación del ingreso, en relación al Índice de Theil. Por tanto, considerando lo anterior, se utilizará el análisis de varianza para descomponer la desigualdad del ingreso, dentro de la Población Económicamente Activa del Área Metropolitana de Monterrey, a través de varios factores componentes. Se utilizará el logaritmo de la varianza ya que se observa que la distribución del logaritmo del ingreso es aproximadamente normal.

2.2 Metodología

Las “funciones generadoras de ingreso” son herramientas estadísticas utilizadas para explicar la desigualdad en el ingreso y otros resultados económicos y de comportamiento. Estas funciones son comúnmente asociadas con un marco de capital humano. Pueden ser vistas como una forma reducida de una ecuación que sintetiza la relación entre el ingreso y varias características personales y regionales.¹⁰ Siguiendo esta interpretación desarrollamos los métodos de regresión y análisis de varianza como funciones generadoras de ingreso, estimándolas empíricamente utilizando datos micro-económicos del Área Metropolitana de Monterrey.

La idea básica del análisis de varianza es descomponer la varianza de la variable dependiente en dos tipos de efectos: aquellos causados por variaciones entre diferentes grupos y los causados por la variación dentro de cada grupo. Por ejemplo, si la variable dependiente es el logaritmo del ingreso de un individuo perteneciente a la P.E.A. y la variable de clasificación o explicativa (factor) es el sector de la actividad ocupacional en la cual esta persona se desenvuelve, la suma de cuadrados total (SCT) del ingreso es expresada como:

$$SCT_y = SC_{entre} + SC_{dentro} \quad (1)$$

donde

$$SCT_y = \sum_i \sum_j (Y_{ij} - \bar{Y})^2 \quad (2)$$

¹⁰ Véase Grilches (1977) y Fields (1982) y (1979).

\bar{Y} es la media total del log del ingreso Y en la muestra entera, las j 's son individuos y las i 's son varias actividades ocupacionales;

$$SC_{\substack{\text{entre} \\ \text{sectores}}} = \sum_i N_i (\bar{Y}_i - \bar{Y})^2 \quad (3)$$

donde \bar{Y}_i es el ingreso medio del sector i y N_i es el número de individuos muestreados en el sector i ; y

$$SC_{\substack{\text{dentro} \\ \text{sectores}}} = \sum_i \sum_j (Y_{ij} - \bar{Y}_i)^2 \quad (4)$$

En otras palabras, la ecuación (1) nos dice la importancia relativa de la desigualdad del ingreso dentro de sectores de actividades ocupacionales comparado con la diversidad en los ingresos medios entre sectores.

Una forma alternativa de ver al modelo es la siguiente (para un solo factor o variable explicativa):

$$Y_{ij} = \mu_{\bullet} + \tau_i + \varepsilon_{ij}; \quad (5)$$

donde: $\mu_{\bullet} = \sum_{i=1}^r \mu_i / r$ es una componente constante (la media total) para todas las observaciones;

μ_i es la media de los valores de respuesta de las observaciones en el i -ésimo nivel del factor bajo estudio;

$\tau_i = \mu_i - \mu_{\bullet}$ es el efecto del i -ésimo nivel del factor a estudiar (o división de la variable explicativa) constante para cada nivel del propio factor;

ε_{ij} son los términos de error, independientes $N(0, \sigma^2)$. $i=1, \dots, r$; $j=1, \dots, n_i$;

r niveles en el factor (o subdivisiones de la variable explicativa); n_i individuos en el i -ésimo nivel.

El modelo ANOVA no requiere supuestos sobre la naturaleza de la relación estadística entre la variable de respuesta y las explicativas, ni que las variables explicativas sean cuantitativas. Tampoco importa el orden en que sean incorporadas al modelo las variables explicativas, por el tipo de método adoptado más adelante para calcular las

sumas de cuadrados de los factores componentes. Asimismo se pueden estudiar las interacciones que se producen entre las diferentes variables explicativas.

Una de las mayores ventajas de las técnicas del análisis de varianza es que indica la importancia cuantitativa de cada variable explicativa. Luego entonces podemos aprender de las descomposiciones por el análisis de varianza cuánta diferencia (en porcentajes por categoría explicativa) se produce, si una persona labora en un sector económico en lugar de otro o si tiene un año adicional de estudios. Ninguno de los 4 procedimientos de descomposición vistos nos da tales resultados.¹¹

¹¹ Una excelente explicación y ejemplificación de las diferencias y similitudes entre el ANOVA y el análisis de regresión, así como de los supuestos y propiedades del ANOVA, se encuentra en Neter, Kutner, Nachtsheim y Wasserman (1996, Cap. 16).

Capítulo III. Variación del Ingreso en el Área Metropolitana de Monterrey, Un Modelo Lineal

3.1 Los datos

Las observaciones utilizadas en el presente trabajo, como se mencionó anteriormente, se obtuvieron de la Encuesta de Migración, Empleo y Salarios en el Área Metropolitana de Monterrey, Marzo-Abril de 1996, levantada por el Centro de Investigaciones Económicas de la Universidad Autónoma de Nuevo León, la cual al utilizar técnicas de muestreo apropiadas nos proporciona una muestra representativa de la población del Área Metropolitana de Monterrey.

Se seleccionaron aquellos individuos que pertenecían a la población económicamente activa (PEA) y recibían un ingreso, bajo los siguientes criterios: 1. Trabajó la semana pasada; 2. No trabajó, pero tenía trabajo; 3. Buscó trabajo. Cabe señalar que después de las depuraciones hechas la gente que *trabajó* constituyó el 96.4% de los casos analizados.

Para determinar el ingreso, en la encuesta se preguntó “¿cuál es su ingreso en pesos? y señale si es mensual, quincenal o semanal, así como si lo obtiene antes o después de impuestos.” Se eliminaron aquellas observaciones en donde no se especificó el periodo, ni el monto del ingreso, ni si lo recibió antes o después de impuestos. Posteriormente se estandarizaron los ingresos a aquellos recibidos semanalmente y después de impuestos, para ulteriormente utilizar el salario por hora recibido la semana anterior a la encuesta y de esta forma hacer más comparables los ingresos recibidos entre los miembros de la muestra.¹²

Asimismo es importante señalar que se trabajó a lo largo de esta Tesis con el ingreso total de las personas (sueldos, salarios, bonos) para descomponer la desigualdad

¹² Por ejemplo, utilizar el salario por hora es la medida más adecuada en este caso para hacer comparaciones entre personas que laboran en distintas actividades, no sería correcto comparar un ingreso semanal de \$40,000 de una persona que trabajó 40 Hrs., con uno de \$20,000 de otra que trabajó sólo 5 Hrs.

entre las mismas, dado que varios estudios han demostrado que en el caso mexicano la proporción más alta del ingreso total la constituyen los ingresos por sueldos y salarios, seguida por la de la propiedad y por último la de las transferencias.¹³ Además, la desigualdad en el ingreso por sueldos y salarios ha explicado la mayor parte de la desigualdad total. Cabe recalcar que dada la naturaleza de la captura de los datos el ingreso manejado en la presente Tesis, es el *ingreso total* que una persona recibió la semana previa a la encuesta, es decir lo que ésta percibió como ingreso, por lo que se puede incluir, además de los salarios y sueldos, algunas otras percepciones como bonos por productividad, para la despensa, la gasolina, etc.

En el estudio se utilizará el logaritmo natural del ingreso (por hora) el cual tiene una distribución aproximada a la “normal” y de esta forma cumplimos con el supuesto de normalidad que requiere el análisis de varianza.

3.2 Exploración y análisis de la base de datos

En el cuadro 3.1 (siguiente página) se muestran los ingresos promedios (por hora) de varias características de los individuos. Entre las diferencias más interesantes a grosso modo se encuentran las siguientes: en el renglón educativo, en general se observa que los ingresos se van incrementado a medida que una persona avanza en el grado de estudio alcanzado, al igual que en los grupos de edad manejados. Sin embargo, no hay un incremento constante en los ingresos al considerarse el efecto de la antigüedad en el último empleo, alcanzándose un máximo entre los 15 y 25 años de servicio. Mientras que no se observan mayores diferencias en la categoría de sexo, ni entre las personas sindicadas o no sindicadas. Cabe aclarar que se utilizó tanto la antigüedad como la edad en el modelo, puesto que la antigüedad es medida sólo para el último empleo, no como antigüedad de por vida.

¹³ Véase Gollás (1982) y Schweitzer (1997).

Cuadro 3.1 Ingreso promedio por hora trabajada la semana pasada a la encuesta (Pesos corrientes, 1996)

Sexo		Grupos de Edad		Último Grado de Estudio Alcanzado	
(1) Masculino	14.56 [791]	(1) 10-19 años	4.26 [84]	(1) Sin estudios	5.9 [26]
(2) Femenino	13.83 [274]	(2) 20-24 años	8.36 [179]	(2) Primaria	7.21 [235]
.	.	(3) 25-29 años	14.9 [151]	(3) Secundaria	7.2 [283]
.	.	(4) 30-34 años	12.29 [126]	(4) Comercial	12.65 [50]
.	.	(5) 35-44 años	17.38 [266]	(5) Técnico	7.96 [70]
.	.	(6) 45-54 años	19.77 [180]	(6) Preparatoria	16.91 [161]
.	.	(7) 55-más años	18.63 [79]	(7) Licenciatura	29.87 [224]
.	.	.	.	(8) Postgrado	51.1 [16]

Antigüedad en el Último Empleo		Situación Sindical		Tipo de Empleo	
(1) 0-9 años	11.96 [784]	(1) Si	13.31 [239]	(1) Empleado	13.81 [638]
(2) 10-14 años	24.25 [76]	(2) No	15.29 [718]	(2) Obrero - Jornalero/Peón	5.71 [186]
(3) 15-19 años	20.31 [82]	(3) No sabe	10.62 [108]	(3) Trabajador por su cuenta	12.16 [149]
(4) 20-24 años	18.29 [57]	.	.	(4) Patrón o empresario	40.99 [85]
(5) 25-más años	20.86 [66]	.	.	(5) Trabajador sin pago en el negocio o predio familiar*	19.72 [7]

Estado Civil		Actividad Económica	
(1) Soltero	9.43 [328]	(1)	16.74 [13]
(2) Casado	16.05 [687]	(2)	11.32 [285]
(3) Separado	23.74 [50]	(3)	11.62 [12]
.	.	(4)	10.95 [91]
.	.	(5)	15.42 [184]
.	.	(6)	12.84 [75]
.	.	(7)	35.8 [27]
.	.	(8)	9.32 [32]
.	.	(9)	18.43 [125]
.	.	(10)	33.39 [41]
.	.	(11)	10.11 [41]
.	.	(12)	11.33 [139]

El número de personas por clasificación se muestra entre corchetes

Números entre paréntesis son clasificatorios

Nomenclatura utilizada en Actividad Económica

- 1.- Agricultura, Ganadería, Caza, Silvicultura, Pesca y Servicios Relacionados, Minería y Extracción de Petróleo y Gas Natural.
- 2.- Industria Manufacturera, de Producción de Alimentos, Bebidas, Tabaco y Textiles.
- 3.- Electricidad y Agua.
- 4.- Construcción.
- 5.- Comercio.
- 6.- Transporte y Comunicaciones.
- 7.- Servicios Financieros de Admon., Alquileres de Muebles e Inmuebles.
- 8.- Servicios de Admon. Pública, Defensa y Seguridad.
- 9.- Servicios Comunes y Sociales.
- 10.- Servicios Especializados, Profesionales y Técnicos.
- 11.- Restaurantes y Hoteles.
- 12.- Servicios de Esparcimiento, Recreativos y Deportivos.

*Las personas dentro de esta categoría son aquellas que no reciben pago por nomina, sin embargo son las "cabezas" del negocio, el cual paga impuestos, y por tanto reciben un ingreso promedio mensual, quincenal o semanal.

Es importante mencionar que en el presente estudio se decidió utilizar categorías para clasificar el grado de estudios, porque se observa que las mayores ganancias por estudiar se dan al terminar un grado escolar.¹⁴ Siendo muy importante diferenciar entre aquellos que tienen un grado de Licenciatura y los que poseen estudios de Postgrado, de otra manera los resultados para el nivel más alto de estudios estarían sesgados.¹⁵

Por otra parte las personas que laboran en el Sector de Servicios Financieros de Administración resultaron ser los que mejores ingresos obtienen, mientras que los menos remunerados son los que trabajan en el Sector Servicios de Administración Pública, Defensa y Seguridad., quienes reciben el 26% del ingreso de los mejor remunerados por actividad económica. En la categoría de estado civil se observa que las personas casadas duplican el ingreso de una persona soltera y aquellas separadas reciben los mayores ingresos.

Para estudiar las posibles interrelaciones entre variables se diseñó un grupo de gráficas, las cuales reflejan el comportamiento del ingreso promedio entre los grupos clasificatorios de una variable explicativa, controlando a través de otra variable explicativa (variable de control).

A continuación se presentan dos gráficas que ejemplifican lo dicho.¹⁶ Las líneas representan la *variable de control*, en este caso el sexo, pero la variable dependiente sigue siendo el ingreso promedio de las personas según la clasificación que les corresponde. La media manejada es la geométrica y se forma para cada "celda" que se crea al combinar los diversos niveles de la variable independiente (en el eje horizontal) con aquellos de la variable de control, utilizando el logaritmo natural del ingreso de las personas bajo estudio. Se utilizan líneas únicamente con el fin de observar las tendencias que sigue el ingreso promedio al variar los niveles de un factor (en la gráfica 1 serían los grados educativos) y mantener constante los de otro (en el ejemplo serían la clasificación por sexo), aunque los factores en sí mismos no sean de tipo cuantitativo sino cualitativo.

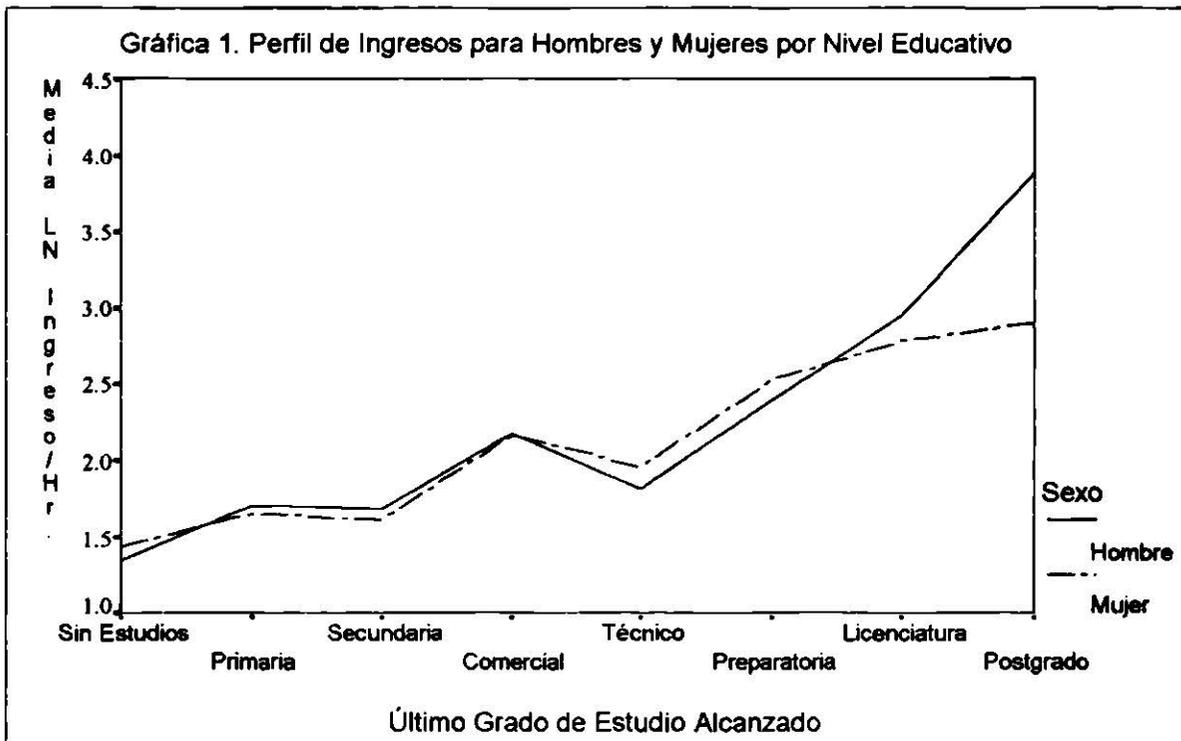
Entre los resultados más interesantes vemos que hay movimientos casi paralelos

¹⁴Véase Meléndez (1995).

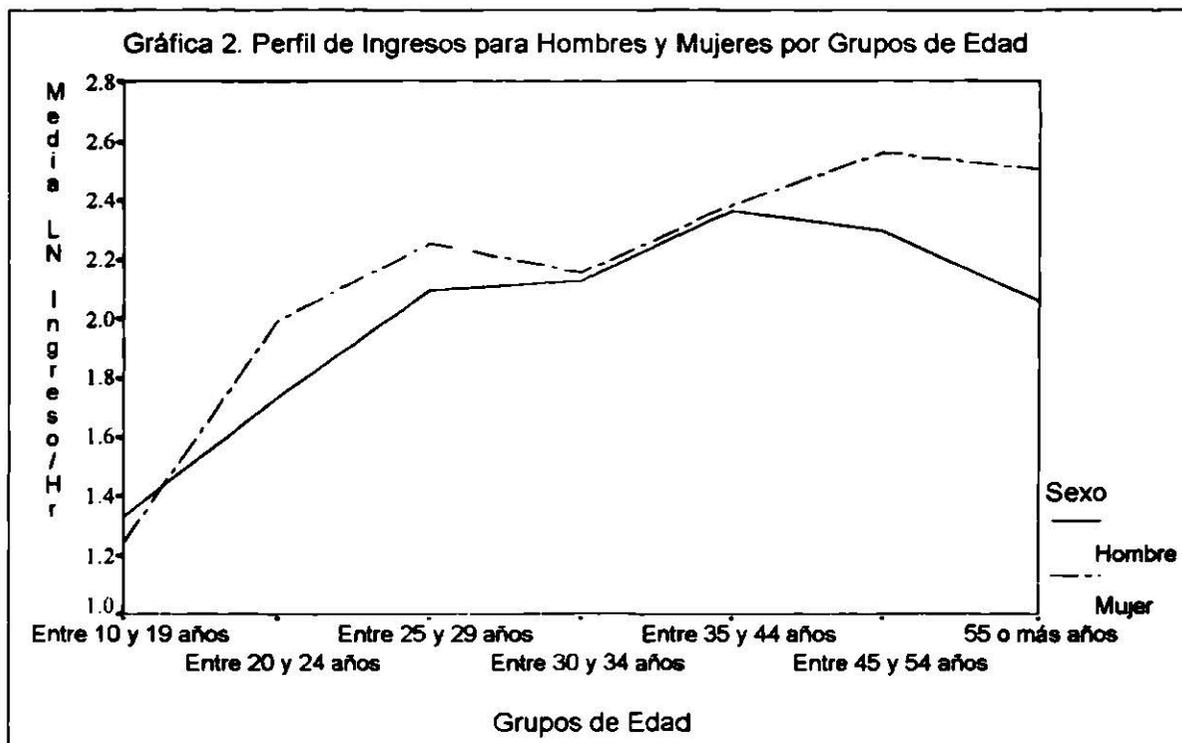
¹⁵ Véase Rupert, Severance-Losin, and Turner (1996).

¹⁶ Una magnífica explicación de cómo interpretar este tipo de gráficas e identificar posibles interacciones se encuentra en Neter, Kutner, Nachtsheim y Wasserman (1996, Cap.19).

entre las tendencias que siguen, a lo largo de los diferentes niveles de factores de las diversas variables, los ingresos controlados por sexo. Es decir, si un hombre gana más al incrementar 5 años su edad, también lo hace una mujer. Esto sugiere una baja interacción entre dichas variables, lo cual se verá con mayor detalle y claridad cuando utilizemos el modelo de análisis de varianza con interacciones.



Así, el eje horizontal de la gráfica 1 presenta los niveles de factores (o categorías) de la variable “último grado de estudio alcanzado” y en la gráfica 2 se presentan los niveles del factor para “edad” (utilizando grupos de edad), para observar las tendencias que siguen los ingresos tanto de los hombres como de las mujeres.



Se observa que, en promedio, las mujeres pertenecientes a la PEA del Area Metropolitana de Monterrey ganan más por hora, *en varios de los niveles estudiados*, que los hombres. Tanto en las categorías de antigüedad en el último empleo, como en algunos niveles de educación. Solamente por nivel educativo se observa claramente que el hombre obtiene un mayor ingreso en el nivel de postgrado, en promedio.¹⁷

Hay que tener en mente que las mujeres incluidas en el estudio son las pertenecientes a la PEA del AMM y que la muestra manejada arroja una relación de casi 3 hombres por cada mujer por lo que se debe ser cauteloso con las relaciones presentadas en las gráficas. Además, las mujeres consideradas poseen los siguientes 3 grados académicos, por orden de importancia (descendente): licenciatura, secundaria y preparatoria; a su vez los hombres de la muestra poseen: secundaria, primaria y licenciatura, también en orden de importancia. Por tanto, el efecto observado en algunas

¹⁷ Para el lector interesado, se presentan más gráficas en el apéndice I del presente estudio. En éstas se observa un comportamiento semejante al descrito para las variables sexo y antigüedad (y sexo y educación).

gráficas se puede deber a un fenómeno de *autoselección* de las mujeres *más capaces*, que deciden incorporarse al mercado laboral.¹⁸

Para resumir sistemáticamente las comparaciones de los diferenciales en la distribución del ingreso por características personales y por la actividad económica, un marco estadístico es útil. Por este motivo un modelo lineal se adopta en la siguiente sección.

3.3 Cuantificación de los efectos personales

En esta sección se presentan los resultados de estimar un modelo lineal de la determinación del ingreso en el Área Metropolitana de Monterrey.

A continuación se presenta el cuadro 3.2 en donde se resume la información final proporcionada por diversas regresiones con las cuales se analiza la contribución explicativa de cada categoría expuesta en el cuadro 3.1.

Cuadro 3.2 Regresiones por factor explicativo

Número de Regresión	1	2	3	4	5
Nombre de la Variable	Sexo	Situación Sindical	Estado Civil	Antigüedad en el Último Empleo	Grupos de Edad
Número de Variables de Engaño	1	2	2	4	6
<i>R- cuadrada</i>	0.0012	0.0048	0.0397	0.0673	0.1021
Error Estándar	0.9526	0.9513	0.9345	0.9218	0.9053
Valor de la <i>F</i>	1.2304	2.5716	21.9442	19.1078	20.0588
Valor de <i>P</i> de la <i>F</i>	0.2676	0.0769	0.0000	0.0000	0.0000

¹⁸ Fenómeno también observado para el AMM por Camero (1995), donde se señala: “*con respecto a los niveles educativos de hombres y mujeres, es de resaltar el hecho que las mujeres pertenecientes a la PEA estén “más educadas” que los hombres. Así se desprende de los resultados de la muestra, que indican que en promedio tienen más años de escolaridad; además, el porcentaje de ellas laborando en actividades profesionales es mayor que el de los hombres.*”

Cuadro 3.2 Regresiones por factor explicativo, continuación.

Número de Regresión	6	7	8	9
Nombre de la Variable	Actividad Económica	Tipo de Empleo	Último Grado de Estudio Alcanzado	Todas las Variables
Número de Variables de Engaño	11	4	7	37
<i>R- cuadrada</i>	0.1101	0.1387	0.3205	0.4567
Error Estándar	0.9034	0.8858	0.7879	0.7148
Valor de la <i>F</i>	11.8385	42.6746	71.2105	23.337
Valor de <i>P</i> de la <i>F</i>	0.0000	0.0000	0.0000	0.0000

Muestra = 1065 Observaciones.

Utilizando variables de engaño (Dummy) diferenciamos entre las subdivisiones de cada categoría. Por ejemplo en la regresión 1 se utilizó una variable de engaño para distinguir entre hombres y mujeres,¹⁹ siendo la clasificación “hombre” la que se desvía de la clasificación “mujer.” Al utilizar variables de engaño, por construcción de nuestro modelo, evitamos el problema de multicolinealidad entre las variables y el problema de autocorrelación.²⁰

Podemos observar que la regresión 1, utilizando al sexo como variable explicativa, no es estadísticamente significativa o útil (observando el valor de *P* del estadístico “*F*” calculado), es decir, con la evidencia empírica con la que contamos no podemos afirmar que esta variable influya de alguna manera en las divergencias de los ingresos de las personas pertenecientes a la PEA del AMM.

Por otro lado el factor último grado de estudio alcanzado es el que mayor nivel explicativo posee, con un 32.05%, utilizando tanto el criterio de la “*F*” calculada así como el del “Error Estándar”. Seguido por el tipo de empleo (con un 13.87%) y en tercer lugar (con un 11.01%) encontramos al factor actividad económica. Este es un punto interesante, dado que una de las variables a la que se le podría atribuir un poder explicativo importante (la actividad económica donde labora una persona) tiene un menor peso en relación con el factor grado de estudio alcanzado, el cual explica por sí mismo el

¹⁹ La variable toma un valor igual a 1 si es hombre y cero si es mujer.

²⁰ Véase Neter, Kutner, Nachtsheim y Wasserman (1996, Cap. 6 y 10).

triple de la variabilidad del ingreso, en relación al primero. A pesar de que en la actividad económica se están estudiando y apreciando variaciones entre 12 sectores.²¹

Los últimos datos que se presentan (regresión # 9) corresponden a la regresión lineal múltiple que engloba a todas las subdivisiones creadas utilizando las variables de engaño. En total se utilizaron 37 variables de engaño Y el porcentaje de la variación del ingreso explicada por la variación de nuestras variables independientes en esta última regresión es de 45.67%, el cual es un nivel aceptable de explicación si tomamos en cuenta distintos trabajos que sobre el tema se han hecho.²²

3.4 Análisis de varianza, modelo de efectos principales

3.4.1 Introducción al análisis de varianza

El Análisis de Varianza (ANOVA por sus siglas en Inglés) es una herramienta muy utilizada en las Ciencias Sociales. Lo encontramos comúnmente cuando consideramos los modelos de regresión. La función de regresión describe la naturaleza de la relación estadística entre la respuesta media y los niveles de las variables de predicción. El ANOVA se utiliza en el análisis de regresión para una variedad de pruebas concernientes con los coeficientes de la propia regresión, el ajuste del modelo, etc. La utilización del ANOVA es en realidad mucho más general que su uso indicado en las regresiones. Los modelos de ANOVA son un tipo básico de un modelo estadístico. Estos se ocupan, como los modelos de regresión, de la relación estadística entre una o más variables de predicción (o variables explicativas) y una variable de respuesta (o dependiente). Además, como los modelos usuales de regresión, la variable de respuesta para el ANOVA es una variable cuantitativa.

²¹ Existen varios trabajos donde se remarca la importancia del sector económico o el giro de la empresa en la determinación de los ingresos individuales, incluso se les atribuye un poder explicativo mayor que a variables de capital humano. Véase Groshen (1991) y Gatica, Mizala y Romaguera (1995).

²² Véase Fields (1980a) y Fields y Schultz (1982).

Los modelos de análisis de varianza difieren de los modelos ordinarios de regresión en dos aspectos clave:

- 1 Las variables de predicción en los modelos de análisis de varianza pueden ser cualitativas (género, ubicación geográfica, cambio de turno, etc.).
2. Si las variables de predicción son cuantitativas, no se requieren hacer supuestos en los modelos de análisis de varianza sobre la naturaleza de la relación estadística entre éstas y la variable de respuesta. Así, no surge la necesidad de especificar en el análisis de varianza la naturaleza de la función de regresión encontrada en el análisis de regresión ordinario.

Es conveniente enfatizar la terminología comúnmente utilizada en este tipo de análisis, principalmente cuando hablamos de las variables explicativas que integran nuestro modelo. En el análisis de varianza las variables explicativas o predictivas son llamadas *factores* y las divisiones dentro de cada variable explicativa, o factor, se denominan *niveles del factor*.

Los elementos básicos del modelo ANOVA para un estudio de un solo factor son muy simples (y se pueden extender a modelos multifactores). Para cada nivel de un factor corresponde una distribución de probabilidad de las respuestas. El modelo supone lo siguiente:

1. Cada distribución de probabilidad es normal.
2. Cada distribución de probabilidad tiene la misma varianza.
3. Las respuestas para cada nivel del factor son muestras aleatorias de la correspondiente distribución de probabilidad y son independientes de las respuestas de cualquier otro nivel del factor.

En los modelos ANOVA, como en cualquier otro modelo estadístico, no es probable que se logre cumplir con todos los supuestos en una situación de la vida real. Sin embargo, éstos se cumplen aproximadamente en muchos casos. Los procedimientos estadísticos sobre los cuales se basa el modelo ANOVA son muy robustos de tal forma

que aún si las condiciones actuales difieren sustancialmente de aquellos requeridos por el modelo ANOVA, el análisis estadístico será aún una aproximación apropiada.

Los parámetros del modelo de análisis de varianza son normalmente desconocidos y deben ser estimados. Al igual que en los modelos de regresión, el método de mínimos cuadrados y el método de máxima verosimilitud conducen a los mismos estimadores de los parámetros del modelo.

El estadístico de prueba a ser utilizado para probar la significancia de los efectos principales y las interacciones es:

$$F^* = \frac{CMeFactor}{CMeError} \quad (6)$$

Donde el numerador y el denominador representan los cuadrados medios del factor en cuestión y del error, respectivamente.

3.4.2 Aplicación del análisis de varianza al ingreso individual

El análisis de varianza, como hemos visto, puede descomponer la variabilidad del ingreso en sus componentes ligados a la demanda laboral y la oferta de factores. Una forma alternativa de ver el modelo de ANOVA, enlazándolo con las regresiones, se basa en la siguiente ecuación generadora de ingreso, utilizando sólo dos factores como ejemplo:

$$W_{ijk} = \mu + X_i\alpha + Y_j\beta + X_iY_j\gamma + e_{ijk} \quad (7)$$

donde

W_{ijk} = ln(Ingreso por hora) del empleado k con un tipo de empleo i en el sector económico j ,

X_i = vector de variables de engaño del tipo de empleo,

Y_j = vector de variables de engaño del sector económico,

X_iY_j = Interacción entre el tipo de empleo i y el sector económico j , y

e_{ijk} = Un término de error.

Así, la varianza del ingreso puede ser descompuesta de la siguiente manera:

$$\sigma_w = \sigma_\alpha + \sigma_\beta + \sigma_\gamma + \sigma_\epsilon \quad (8)$$

El tamaño de cada componente de varianza indica su importancia económica relativa y puede ser entendido como sigue:

1. *Diferencial por el tipo de empleo*, σ_α , una desviación de los ingresos medios de cada uno de los tipos de empleo, a lo largo de todos los sectores económicos, presumiblemente reflejando la productividad y compensando por dichas diferencias. Su varianza mide la importancia de factores externos de la oferta laboral por tipo de empleo.
2. *Diferencial por el sector económico*, σ_β , la desviación promedio de la media del ingreso en un sector económico en relación a los otros, a lo largo de todos los tipos de empleo. Su varianza mide el impacto completo de las diferencias en el sector económico sobre la determinación del ingreso.
3. *Diferencial por la celda de interacción*, σ_γ , pago a los trabajadores manteniendo el mismo tipo de empleo y el mismo sector económico sobre los diferenciales del tipo de empleo y el sector económico. Dado que este término ajusta por diferencias en la estructura interna de la distribución del ingreso entre los individuos, su varianza mide la independencia de los mercados de trabajo internos.
4. *Diferencial individual dentro de los factores*, σ_ϵ , una desviación individual o residual de la media formada por todos los factores considerados, presumiblemente como resultado de diferencias en la productividad del individuo, o por diferentes estrategias compensadoras de los empleadores. Su varianza captura la importancia de las diferencias individuales.

Dado que los datos están desbalanceados, se crea un problema de no ortogonalidad en el modelo ANOVA por lo que un modelo de análisis de varianza no paramétrico es utilizado. Este método evita imponer estructuras a los diferenciales estimados, lo cual no sería deseable en un ejercicio de investigación como este.²³

²³ El desbalance en los datos se debe a una desigual distribución de las observaciones a lo largo de los diferentes factores estudiados. Las técnicas para la estimación de los componentes de la varianza en un modelo desbalanceado se detallan en Searle (1971).

El problema del desbalance en los datos dificulta el análisis de varianza, al volverse más complejas las fórmulas de las ecuaciones de mínimos cuadrados y no poderse sumar los efectos de los factores (variabilidad explicada) para obtener directamente la suma de cuadrados del total explicado. Sin embargo, utilizando un método alternativo de cálculo de la suma de cuadrados se logra superar el problema. El método empleado fue el de la suma de cuadrados tipo III del paquete estadístico SAS, el cual calcula los efectos de cada factor tomando en cuenta al resto de los factores al mismo tiempo, incluyendo las interacciones, de tal suerte que los efectos principales y por interacción, del análisis de varianza del ejemplo anterior de dos factores, quedarían de la siguiente manera:

Cuadro 3.3 Cálculo de un modelo ANOVA desbalanceado

Efecto	Suma de Cuadrados Tipo III
Tipo de Empleo	$R(\alpha \mu, \beta, \alpha\beta)$
Sector Económico	$R(\beta \mu, \alpha, \alpha\beta)$
Interacción	$R(\alpha\beta \mu, \alpha, \beta)$

Fuente: The SAS System for Linear Models (1991).

La interpretación de la variabilidad explicada por cada factor es igual a la anteriormente presentada, pues los resultados de este tipo de análisis son conceptualmente análogos a los obtenidos en el modelo de análisis de varianza con datos balanceados, esto es, seguimos observando la variabilidad explicada por cada factor y si ésta es, o no, estadísticamente significativa.²⁴

Un modelo sin interacciones se reporta en el cuadro 3.4. Se observan los distintos niveles de significancia que existen entre las categorías analizadas y la proporción de la variabilidad explicada por la categoría correspondiente, la cual es la suma de cuadrados de esa categoría dividida entre la suma de cuadrados total. Cabe señalar que se hicieron

²⁴ Para una explicación más profunda sobre la descomposición de la varianza en modelos desbalanceados, véase Neter, Kutner, Nachtsheim y Wasserman (1996, Cap. 22), y The SAS System for Linear Models (1991).

pruebas para comprobar la validez de los supuestos detrás del ANOVA (por ejemplo, la normalidad en la distribución de los errores, varianzas constantes, no “outliers,” etc.) y se encontró que se cumplía con ellos.²⁵

Cuadro 3.4 ANOVA de efectos principales

Fuente de Variación	Proporción de la Varianza Explicada (%)	"F" Marginal ^c	Valor de "P"
Último Grado de Estudio Alcanzado	15.908	42.96	0.0001
Tipo de Empleo	2.610	12.34	0.0001
Grupos de Edad	1.707	5.38	0.0001
Actividad Económica	1.312	2.26	0.0104
Antigüedad en el Último Empleo	0.968	4.58	0.0011
Estado civil	0.505	4.77	0.0086
Situación Sindical	0.063	0.60	0.5514
Sexo	0.013	0.25	0.6194
Covarianza	22.588		
Total Explicado ^a	45.674	23.34	0.0001
Individual ^b	54.326		
Total	100		

Muestra = 1065 Observaciones.

- a. Se refiere a la variación total explicada entre los diversos niveles de los factores.
- b. Variabilidad dentro de los niveles de los factores.
- c. Marginal, por ser calculada para cada factor explicativo y para el total explicado.

Nuestros resultados son consistentes con los encontrados en el modelo de regresión lineal. Nuevamente la categoría educación posee, por mucho, el mayor peso explicativo con un 15.9%, mientras que le siguen el tipo de empleo con un 2.61%, la edad con un 1.7%, la actividad económica en donde se labora con un 1.31%, la antigüedad en el último empleo con un 0.97% y el estado civil con un 0.51%. Las variables sexo y situación sindical no poseen un efecto estadísticamente significativo. Es decir, casi una quinta parte de los diferenciales en ingresos puede ser explicada únicamente por el grado de estudios alcanzado por una persona, corroborando algunas hipótesis planteadas en los estudios de capital humano.

²⁵ En el apéndice 2 se muestran los resultados de las pruebas estadísticas, con las cuales se verificó la validez de los supuestos.

Es interesante observar, utilizando el criterio de la F calculada, que incluso los grupos de edad poseen un mayor poder explicativo que la categoría de actividad económica. Por tanto, cuando investigamos las causas de la variabilidad del ingreso por factores, para la población económicamente activa del Área Metropolitana de Monterrey, el grado de educación alcanzado es el factor explicativo más importante y la actividad económica, contrario a lo que se podría esperar, parece no reflejar diferencias sustanciales entre los trabajadores de las diversas ramas económicas, una vez que controlamos por variables de capital humano. Lo cual podría deberse a una baja concentración industrial que evita “privilegios” para la mayoría de los trabajadores de un sector en relación con el resto.

Así, el mercado laboral del AMM presenta signos de ser *competitivo*, en el sentido de remunerar mejor a quienes poseen mayor capital humano.²⁶ Además, no se observan grandes variaciones del ingreso entre personas con el mismo capital humano laborando en distintos sectores económicos, ni se presentan diferencias significativas entre hombres y mujeres, así como tampoco entre sindicatos y no sindicatos. Controlándose estas diferencias por el resto de los factores.

Es importante señalar que la variabilidad explicada por cada factor integrante del modelo ANOVA precedente, difiere de aquella presentada por cada factor en los análisis de regresión, debido a que en el último caso se evalúa la variabilidad del ingreso en función de una sola variable (o factor), mientras que en el modelo ANOVA permitimos la incorporación de varios factores. Nótese que al incorporarse al modelo todos los factores explicativos, capturando la variabilidad de los efectos principales (es decir, sin interacciones), se genera un factor de *covariabilidad*, el cual captura las variaciones del ingreso entre todos los niveles de los factores del modelo simultáneamente (38 en este caso). La covariabilidad puede conceptualizarse como la diferencia entre la variabilidad explicada por cada factor independiente (en las regresiones, por ejemplo) y su contribución marginal dentro del modelo ANOVA completo. Este componente es resultado del método no paramétrico utilizado.

²⁶ Sabemos que se incrementan los ingresos al ascender en el escalafón educativo por medio de los datos presentados en el cuadro 3.1 y el análisis de los mismos en la sección 3.2, además del análisis gráfico (gráficas 1 y 2). El lector interesado puede consultar más gráficas en el apéndice 1 del presente trabajo.

Sin embargo, dado que se utilizó el mismo número de subdivisiones en cada categoría para obtener la suma de cuadrados totales tanto de la regresión múltiple, que se mostró en la sección anterior, como en la del total explicado, que se obtuvo en el análisis de varianza, el porcentaje de variación explicado (R^2 , en el modelo de regresión) es igual en ambos modelos, 45.67%

Este último resultado nos indica que el 45.67% de la variación del ingreso es explicado por la variación que existe entre los diversos niveles de los factores que integran cada categoría explicativa (por ejemplo las 12 ramas de actividad económica forman los 12 grupos de esta categoría, en donde los grados de libertad son igual a 11, $(r-1)$) y la variabilidad capturada por la covarianza. Es decir, el 45.67% de la variación del ingreso se explica por la variación entre los 38 grupos que conforman el modelo completo. Por lo que un poco más del 50% de la variación del ingreso entre los individuos del Área Metropolitana de Monterrey es explicada por variaciones dentro de cada grupo (variación individual).

La variación individual, como se mencionó, refleja las diferencias propias de los individuos como la productividad, las habilidades innatas, el entorno familiar, las condiciones de salud, raza, etc., así como diferencias en las características de los empleadores, tales como implementar distintas estrategias compensadoras para sus empleados.

Lo cual podría significar que, aún y cuando se controla por características de capital humano, personales y del sector donde se labora, existe una alta diferenciación entre los ingresos que perciben los habitantes del AMM. De tal suerte, la evidencia empírica sugiere que si bien es muy importante el factor educativo en la determinación de los diferenciales del ingreso, existen varios factores adicionales, como el tipo de empresa en la que se labore, las características personales y de los empleadores, que influyen en las diferencias individuales de los ingresos en el Área Urbana de Monterrey. Resultado encontrado en estudios para países sudamericanos,²⁷ y que abre un importante campo de investigación futura.

²⁷ Véase Gatica, Mizala y Romaguera (1995).

3.4.3 Análisis de varianza dentro de grupos educativos

El resultado principal de la sección precedente fue la confirmación de la gran importancia relativa que posee la *educación* en la explicación global de las desigualdades del ingreso. Casi una quinta parte de la variabilidad del ingreso puede ser explicada por ese factor solamente. Por tanto, se investigó dentro de 4 grupos educativos los factores explicativos de la alta variabilidad interna. Los resultados se presentan a continuación.

Cuadro 3.5 ANOVA por nivel educativo

Fuente de Variación	<u>Estudios Superiores</u>			<u>Preparatoria</u>		
	Proporción de la Varianza Explicada (%)	"F" Marginal ^c	Valor de "P"	Proporción de la Varianza Explicada (%)	"F" Marginal ^c	Valor de "P"
Actividad Económica	9.596	2.49	0.0040	3.205	1.25	0.2530
Tipo de Empleo	5.094	4.29	0.0020	4.221	3.63	0.0030
Grupos de Edad	5.651	2.73	0.0100	2.672	1.64	0.1240
Antigüedad en el Último Empleo	2.822	1.91	0.0950	7.027	6.04	0.0001
Estado civil	2.312	1.56	0.1730	3.556	3.06	0.0110
Situación Sindical	0.889	1.50	0.2250	2.228	4.79	0.0090
Sexo	0.121	0.41	0.5230	0.003	0.01	0.9100
Covarianza	13.664			20.343		
Total Explicado ^a	40.149	3.66	0.0001	43.256	5.17	0.0001
Individual ^b	59.851			56.744		
Total	100			100		
Muestra		240			281	

Fuente de Variación	<u>Secundaria</u>			<u>Primaria</u>		
	Proporción de la Varianza Explicada (%)	"F" Marginal ^c	Valor de "P"	Proporción de la Varianza Explicada (%)	"F" Marginal ^c	Valor de "P"
Actividad Económica	3.072	0.93	0.5090	5.900	1.42	0.1580
Tipo de Empleo	2.970	1.99	0.0810	9.700	5.54	0.0001
Grupos de Edad	6.425	3.58	0.0020	2.200	0.89	0.5160
Antigüedad en el Último Empleo	0.314	0.21	0.9580	2.300	1.13	0.3450
Estado civil	2.028	1.36	0.2420	5.800	3.30	0.0070
Situación Sindical	2.330	3.89	0.0220	0.100	0.20	0.8220
Sexo	0.119	0.40	0.5300	0.005	0.02	0.9020
Covarianza	8.808			5.417		
Total Explicado ^a	26.066	2.48	0.0001	31.422	2.36	0.0001
Individual ^b	73.934			68.578		
Total	100			100		
Muestra		283			235	

a. Se refiere a la variación total explicada entre los diversos niveles de los factores.

b. Variabilidad dentro de los niveles de los factores.

c. Marginal, por ser calculada para cada factor explicativo y para el total explicado.

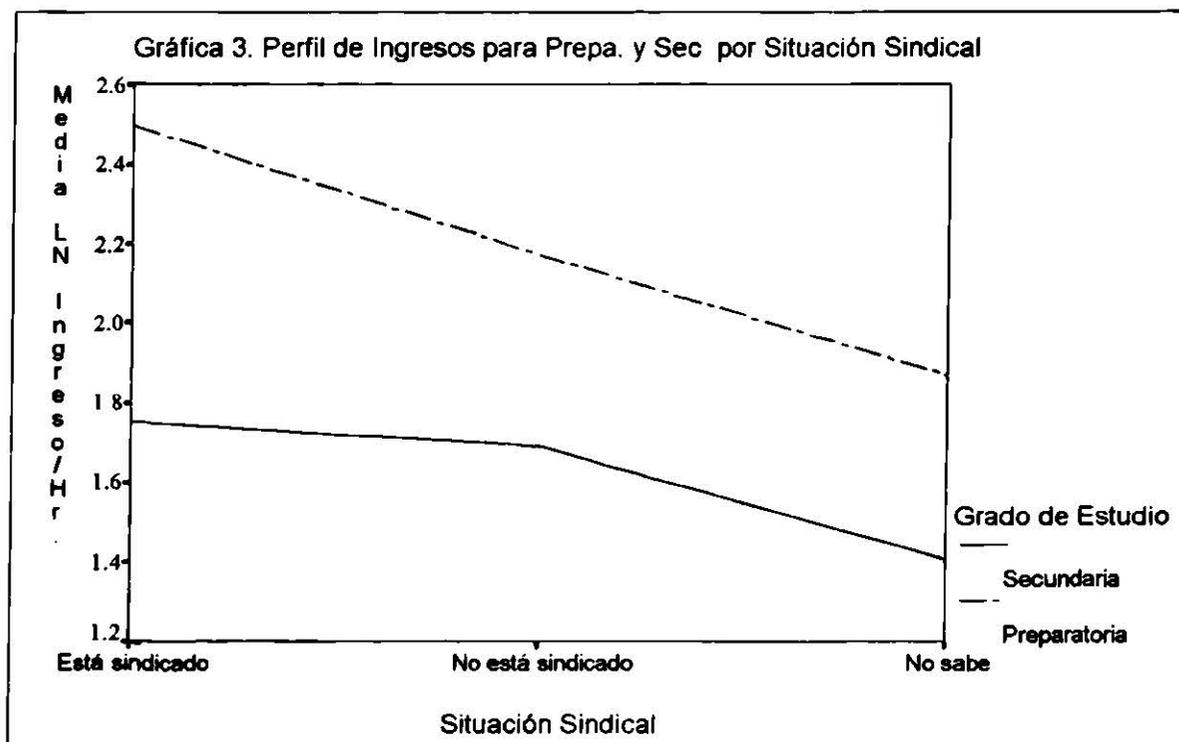
Dentro del grupo de personas con educación superior terminada se observa una alta variabilidad explicada por los 7 factores considerados, un 40.14%, siendo estadísticamente no significativos los factores sexo y situación sindical (ambos como en el modelo ANOVA de efectos principales), así como el estado civil de la persona. Los otros 4 factores arrojan resultados interesantes, siendo el sector de la actividad económica donde labora una persona el factor de mayor importancia en la explicación de la variabilidad del ingreso, con un 9.59%. Algo interesante desde el punto de vista de la integración de los mercados de trabajo, segmentados por nivel educativo, puesto que el factor actividad económica únicamente es significativo para las personas con un nivel superior de estudios, en los demás grupos (medio superior, secundaria y primaria) no figura como un factor explicativo válido. Es decir, la evidencia sugiere que en el mercado laboral del AMM se *premia* a la educación superior de manera marcada en algunos sectores de la actividad económica, con relación al resto de los grados educativos. Para las personas con el grado educativo más alto es muy importante elegir el sector económico al cual se incorporarán, puesto que las diferencias en las percepciones económicas serán amplias y la movilidad laboral, dadas las propias divergencias en ingresos, parece baja. Así, el mercado laboral de las personas con educación superior es el único en el cual la actividad económica juega un papel fundamental, en los otros segmentos educativos parece no influir de manera importante.

Luego entonces, para un profesional de alguna área del conocimiento es muy importante incorporarse al segmento de mercado, en este caso por actividad económica, que mejor se adapte a las características de su profesión. En otras palabras, la demanda laboral de los profesionistas en el AMM parece distinguir claramente entre carreras profesionales, mientras que en el caso de los demás niveles educativos parecería no haber mayores divergencias en ese aspecto.²⁸ Por su parte, los grupos de edad son el segundo factor explicativo más importante para las personas con educación superior, con un 5.65% de la variabilidad explicada, seguido por el tipo de empleo con un 5.09% y la antigüedad en el último empleo con un 2.82%.

²⁸ Los trabajadores se consideran igualmente productivos en los diversos sectores económicos dentro de cada grupo educativo.

Para las personas con educación media superior la antigüedad en el último empleo es el principal factor explicativo con 7.02%, seguida por su tipo de empleo con 4.22%, su estado civil con 3.55% y su situación sindical con 2.22%. Para las personas con educación secundaria la edad explica 6.43%, seguida por su tipo de empleo con 2.97% y por su situación sindical con 2.32%. A su vez, para aquellos con educación primaria la variabilidad del ingreso entre las mismas es explicada solamente por el tipo de empleo que logren conseguir con un 9.7% y por su estado civil con un 5.8%.

De aquí se desprenden 2 puntos: primero, la situación sindical que enfrentan tanto las personas del nivel medio superior así como aquellos con secundaria completa, parece influir significativamente en sus diferenciales de ingreso. Estos segmentos de la PEA podrían enfrentarse a "barreras" impuestas por los sindicatos para pertenecer a los distintos gremios y, según la gráfica 3, en ambos casos conviene pertenecer al sindicato.



Situación que parecen no vivir aquellos con educación superior o primaria. Más adelante se analizarán estos factores con un modelo de ANOVA con interacciones.

En segundo lugar, se observa que para las personas con educación primaria existen sólo 2 factores que resultan válidos en la explicación de la variabilidad de sus ingresos, el estado civil y el tipo de empleo. Esto sugiere que para aquellos con educación primaria la influencia del resto de los factores estudiados es poco significativa en la determinación y dispersión de sus ingresos, es decir, la movilidad laboral de estas personas con relación a la actividad económica y su situación sindical es alta.

Es importante señalar que cada uno de los modelos creados para cada grupo de educación fue significativo en su conjunto. Además de recalcar la no-existencia de diferencias en ingresos entre hombres y mujeres en ninguno de los grupos educativos, sin embargo se deben tener presentes los comentarios sobre la base de datos presentados en el capítulo precedente y el hecho de que el AMM disfrute de un grado de avance económico superior al resto de las zonas urbanas del país. En el siguiente capítulo se analizan a detalle las posibles interacciones sugeridas hasta aquí.

Capítulo IV. Variación del Ingreso en el Área Metropolitana de Monterrey, Un Modelo con Interacciones

En el capítulo anterior se estudiaron los distintos factores que influyen en las desigualdades del ingreso de la Población Económicamente Activa del Área Metropolitana de Monterrey, utilizando tanto un modelo de regresión lineal múltiple, así como el modelo de análisis de varianza de efectos principales. En el presente capítulo se extiende el modelo ANOVA para permitir la incorporación de los elementos de interacción, los cuales mejoran la exactitud del modelo en varios sentidos: incrementan la eficiencia del modelo ANOVA al capturar un mayor porcentaje de la variabilidad del ingreso y pueden fortalecer la validez de los resultados obtenidos en el modelo de efectos principales²⁹ (es decir, sin interacciones).

4.1 Interpretación de las interacciones en el modelo

Como un ejemplo ilustrativo del modelo a utilizar con 8 variables explicativas, se presenta el siguiente modelo general de análisis de varianza con sólo 2 variables explicativas (factores A y B):

$$Y_{ijk} = \mu_{..} + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk} \quad (9)$$

donde: $\mu_{..} = \frac{\sum_i \sum_j \mu_{ij}}{ab}$ es la media total constante.

$\alpha_i = (\mu_{i.} - \mu_{..})$ es el efecto principal por el factor A al i -ésimo nivel. El cual es constante sujeto a la restricción: $\sum \alpha_i = 0$.

$\beta_j = (\mu_{.j} - \mu_{..})$ es el efecto principal por el factor B al j -ésimo nivel. El cual es constante sujeto a la restricción: $\sum \beta_j = 0$.

²⁹ Sin embargo, no siempre sucede esto, por lo que en algunos estudios una vez realizadas las interacciones en el modelo se decide entre utilizar el modelo completo (si la ganancia marginal es significativa) o mantener sólo el de efectos principales, sobre la base de los objetivos de la investigación.

$(\alpha\beta)_{ij} = \mu_{ij} - (\mu_{..} + \alpha_i + \beta_j)$ es el efecto de la interacción cuando el factor A está en el i -ésimo nivel y el factor B en el j -ésimo. Es constante, sujeto a las restricciones:

$$\sum_i (\alpha\beta)_{ij} = 0 \quad j=1, \dots, b$$

$$\sum_j (\alpha\beta)_{ij} = 0 \quad i=1, \dots, a$$

ε_{ijk} son los términos de error, independientes con $N(0, \sigma^2)$ $i=1, \dots, a$ $j=1, \dots, b$ $k=1, \dots, n$; a niveles del Factor A; b niveles del Factor B; n observaciones por celda creada.

Podemos estudiar la existencia de los efectos de interacción del factor formalmente, examinando si las medias de los niveles de cada factor μ_{ij} pueden ser expresadas de la siguiente manera: $\mu_{ij} = \mu_{..} + \alpha_i + \beta_j$, si esto es posible, los efectos principales de los factores explicativos son aditivos; de otra forma, los factores estarán interactuando.

La diferencia entre la media del nivel de un factor μ_{ij} y el valor $\mu_{..} + \alpha_i + \beta_j$, el cual se esperaría si los dos factores fueran aditivos, es llamada el efecto de la interacción, o más sencillo la *interacción*, del i -ésimo nivel del factor A con el j -ésimo nivel del factor B y se denota por $(\alpha\beta)_{ij}$. Por tanto, se define al factor $(\alpha\beta)_{ij}$ de la siguiente forma:

$$(\alpha\beta)_{ij} = \mu_{ij} - (\mu_{..} + \alpha_i + \beta_j) \quad (10)$$

El modelo ANOVA con interacciones, representado por la ecuación (9), se sujeta a los principios y supuestos generales del modelo ANOVA de efectos principales. Además, podemos representarlo a través de una regresión múltiple, al igual que en el capítulo III ecuación (7), la cual se repite aquí.

$$W_{ijk} = \mu + X_i \alpha + Y_j \beta + X_i Y_j \gamma + e_{ijk}$$

donde

$W_{ijk} = \ln(\text{Ingreso por hora})$ del empleado k con un tipo de empleo i en el sector económico j ,

$X_i =$ vector de variables de engaño del tipo de empleo,

$Y_j =$ vector de variables de engaño del sector económico,

$X_i Y_j =$ Interacción entre el tipo de empleo i y el sector económico j , y

e_{ijk} = Un término de error.

La varianza del ingreso puede ser descompuesta de la siguiente manera:

$$\sigma_w = \sigma_\alpha + \sigma_\beta + \sigma_\gamma + \sigma_\epsilon$$

La interpretación de cada elemento es igual a la descrita en el capítulo III, por lo que sólo enfatizaremos la parte concerniente a la interacción:

Diferencial por la celda de interacción, σ_γ , el pago a los trabajadores manteniendo el mismo tipo de empleo y el mismo sector económico sobre los diferenciales del tipo de empleo y el sector económico. Dado que este término ajusta por diferencias en la estructura interna de la distribución del ingreso entre los individuos, su varianza mide la independencia de los mercados de trabajo internos. Es decir, el factor de interacción entre el sector económico y el tipo de empleo compara las diferencias que existen entre los ingresos promedio de los tipos de empleos en un sector económico con sus similares en otro sector, y así a través de todos los sectores. Si la interacción resulta no significativa querría decir que la misma estructura de diferenciales de ingreso en un sector de actividad económica se mantiene a lo largo de las demás actividades económicas.³⁰

4.2 Análisis de varianza, modelo completo con interacciones

Una vez analizados tanto la definición de las interacciones como su incorporación al modelo de ANOVA, se procede a estimar el modelo completo para probar estadísticamente todas las posibles interacciones señaladas con antelación y su interpretación económica.

En el cuadro 4.1 (siguiente página) se presenta el modelo completo³¹ de análisis de varianza. Lo primero que debemos hacer notar es el fortalecimiento del modelo con la incorporación de los términos de interacción. Esto es, el poder explicativo del modelo de

³⁰ Hay que tener en mente, que al revisarse una interacción se están tomando en cuenta las dos "direcciones" del análisis. Por ejemplo, en el esquema presentado para aclarar el concepto de interacción entre el tipo de empleo y la actividad económica, la misma interacción nos da la información de la estructura de salarios (diferenciales de ingreso promedio) entre las ramas de la actividad económica del AMM, para un tipo de empleo, y probar si esta estructura se mantiene a través de los demás tipos de empleo o no. Si se mantiene, entonces no hay interacción entre estas variables.

análisis de varianza pasa de un 45.67% (en el caso del modelo de efectos principales), a un 75.57%, es decir, casi un 30% adicional de la variabilidad del ingreso entre la PEA del AMM capturada por el modelo completo con interacciones. El resultado no es de menor importancia, dado que en diversos estudios realizados para países desarrollados, y algunos sudamericanos, al incorporarse los términos de interacción el modelo final sólo mejora unos cuantos puntos porcentuales su poder explicativo, por lo que se decide abandonar las interacciones como parte del estudio mismo.³²

Puesto que en este caso las interacciones parecen constituir una parte esencial en la explicación de los diferenciales del ingreso de las personas, se explorarán de manera esquemática, todos los efectos capturados en el cuadro 4.1. Cabe aclarar que la proporción de la variabilidad directamente explicada por cada factor estudiado disminuye en relación al porcentaje presentado en el modelo de análisis de varianza de efectos principales, cuadro 3.4. En parte debido al método no-paramétrico utilizado, lo cual se señala en la sección 3.4 donde se presenta una introducción al análisis de varianza y una aplicación al ingreso individual, y en parte por los elementos de interacción con los cuales cuenta ahora el propio modelo de análisis de varianza. Además, otra parte de la variabilidad de cada factor principal es ahora capturada por un efecto originado por el mismo método no-paramétrico, comentado anteriormente, la *covariabilidad* o *efecto conjunto*.

La covariabilidad captura las variaciones del ingreso entre todos los niveles de los factores estudiados, incluyendo las interacciones, simultáneamente. Es decir, la covariabilidad captura el efecto sobre los diferenciales en los ingresos dentro de todos los factores de manera conjunta. Se puede conceptualizar como la diferencia entre la variabilidad explicada por cada factor independiente (visto como la variabilidad capturada por un solo factor en un análisis de varianza individual) y su contribución marginal dentro del modelo de análisis de varianza completo. Más que pensar en ella como una correlación entre las variables explicativas, debería entenderse como el poder explicativo “agregado” de todos los niveles de los factores estudiados, visto como un solo factor.

³¹ El modelo se corrió utilizando la suma de cuadrados del tipo III del paquete estadístico SAS, por las razones expuestas en la sección 3.4.2 del presente estudio.

³² Véase Fields y Schultz (1982) y Mizala y Romaguera (1998).

Cuadro 4.1 ANOVA con interacciones

Fuente de Variación	Proporción de la Varianza Explicada (%)	"F" Marginal^c	Valor de "P"
Último Grado de Estudio Alcanzado (Educación)	1.966*	6.58	0.0001
Actividad Económica (Act.Ecn.)	0.757*	1.61	0.0913
Estado Civil (E.Civil)	0.444*	5.20	0.0058
Tipo de Empleo (T.Empleo)	0.786*	4.60	0.0012
Antigüedad en el Último Empleo (Antigüedad)	0.515*	3.02	0.0177
Sexo	0.063	1.48	0.2242
Situación Sindical (Sit.Sind.)	0.065	0.77	0.4652
Grupos de Edad (Edad)	0.434	1.69	0.1201
Educación*Act.Ecn.	2.508	1.20	0.1733
Educación*E.Civil	0.357	0.84	0.5929
Educación*Antigüedad	1.268	1.29	0.1653
Educación*Sexo	0.224	0.88	0.5124
Educación*Edad	1.645	1.17	0.2420
Educación*Sit.Sind.	1.122*	2.19	0.0111
Educación*T.Empleo	1.240*	1.61	0.0521
Act.Ecn.*E.Civil	0.552	0.86	0.6080
Act.Ecn.*Sit.Sind.	0.642	0.84	0.6590
Act.Ecn.*Sexo	0.596	1.55	0.1272
Act.Ecn.*Edad	2.518	1.16	0.2199
Act.Ecn.*T.Empleo	2.355*	2.40	0.0003
Act.Ecn.*Antigüedad	2.050*	1.45	0.0508
E.Civil*Edad	0.657	1.40	0.1694
E.Civil*Sit.Sind.	0.236	1.38	0.2397
E.Civil*Sexo	0.019	0.22	0.8043
E.Civil*T.Empleo	0.803*	3.13	0.0050
E.Civil*Antigüedad	0.647*	2.16	0.0358
T.Empleo*Antigüedad	0.744	1.45	0.1378
T.Empleo*Edad	0.665	0.92	0.5549
T.Empleo*Sexo	0.043	0.33	0.8010
T.Empleo*Sit.Sind.	0.386	1.51	0.1734
Antigüedad*Edad	0.687	1.15	0.3114
Antigüedad*Sexo	0.157	0.92	0.4535
Antigüedad*Sit.Sind.	0.411	1.20	0.2939
Sexo*Edad	0.162	0.63	0.7035
Sexo*Sit.Sind.	0.021	0.24	0.7839
Edad*Sit.Sind.	0.426	0.83	0.6184
Covarianza	47.40		
Total Explicado^a	75.571*	3.60	0.0001
Individual^b	24.429		
Total	100		

Muestra = 1065 Observaciones.

*Estadísticamente significativas con un coeficiente de confianza del 90%.

- Se refiere a la variación total explicada entre los diversos niveles de los factores.
- Variabilidad dentro de los niveles de los factores.
- Marginal, por ser calculada para cada factor explicativo y para el total explicado.

De esta forma, la disminución en la variabilidad directamente explicada por cada factor principal no desaparece, simplemente se diversifica entre su aportación directa, la covariabilidad y las interacciones de este factor con cada una de las otras variables explicativas.

Así mismo, al ser el estudio de las interacciones el objetivo principal del presente capítulo, es conveniente volver a definir las de manera general, para después analizar caso por caso. Las interacciones presentadas se denominan de “dos vías,” es decir, por pares de factores explicativos. Están diseñadas para probar si se mantiene, o no, la estructura de los diferenciales de los ingresos entre los diversos niveles de un factor, observados “dentro” de un nivel (o subdivisión) de un segundo factor explicativo, cuando se examinan todos los niveles del segundo factor.³³

Así, se puede estudiar la estructura de ingresos en un factor, dentro de un nivel de otro, investigando si ésta cambia a lo largo de las subdivisiones del segundo factor. Con lo cual se probaría estadísticamente si los diferenciales, o no-diferenciales, entre diversas categorías clasificativas se mantienen en todo momento.

Ahora, se inicia el estudio esquemático del cuadro 4.1, para cada uno de los factores analizados.

1. La educación; indiscutiblemente vuelve a ser el factor con mayor proporción de la variabilidad del ingreso explicada de forma directa y estadísticamente hablando es el factor cuyo efecto principal es el más significativo.

Analizando las interacciones entre este factor y los demás, examinaremos si los diferenciales en ingresos (observados o no) en los otros factores, es decir su estructura de ingresos, se mantienen a través de cada uno de los niveles educativos. La información nos ayudará a comprender mejor la posible influencia de los diversos grados educativos en las estructuras de ingresos de los otros factores. Aunque las interacciones en sí mismas no nos dicen explícitamente en cuál nivel educativo cambian los diferenciales en ingresos de

³³ Suponiendo que se hayan apreciado diferencias en los ingresos promedio entre los niveles del primer factor, con el cual se examina una posible interacción con un segundo factor. Esto se puede verificar en el modelo ANOVA de efectos principales, y para observar la dirección de las diferencias (por ejemplo, si se incrementa el ingreso al cambiarse de un sector económico a otro) se puede consultar el cuadro 3.1 ó se pueden consultar las gráficas del apéndice 1.

algún otro factor, si nos dan la pauta para iniciar un análisis más profundo entre ese par de factores, si el investigador así lo considera necesario.

En este caso, el factor “último grado de estudio alcanzado” no interactúa con varios de sus similares: actividad económica, estado civil, antigüedad en el último empleo, sexo ni con los grupos de edad. Esto es, los diferenciales en ingresos promedio entre las diversas actividades económicas se mantienen estadísticamente sin cambios a través de todos los grados educativos.³⁴ Lo cual se puede interpretar de manera inversa, las diferencias en ingresos entre los niveles educativos (premios salariales por grado alcanzado) se mantienen a través de todos los sectores económicos presentados. Resultado que confirma lo sugerido en la sección 3.4 del presente estudio. La importancia del capital humano a lo largo de todas las actividades económicas es reconocida y premiada de manera muy similar.

Con el estado civil sucede lo mismo, las diferencias en ingresos entre las clasificaciones del propio estado civil se mantienen a través de los grados educativos. Así sucede también con los ingresos promedio según la antigüedad en el último empleo, las diferencias se mantienen en cada uno de los niveles educativos. Al igual que con las diferencias en ingresos promedio entre los grupos de edad, la estructura de ingresos no cambia entre los grados educativos. Mientras que las *no-diferencias* en ingresos entre hombres y mujeres parecen no variar a través de todos y cada uno de los niveles educativos presentados, hecho interesante puesto que refuerza nuestros hallazgos previos sobre la posible no-segregación salarial en contra de la mujer, en cuanto a los ingresos por hora se refiere. Es decir, los premios salariales, tanto para hombres como para mujeres, son equivalentes al pasar de un grado educativo a otro.

Asimismo, el factor educativo parece interactuar de forma importante con los factores: situación sindical y con el tipo de empleo. En el último caso, la interacción sugiere cambios en los diferenciales en ingresos entre los diversos empleos a ocupar a lo largo de los grados educativos analizados. Lo cual es hasta cierto punto esperado, puesto que no es factible pensar que las personas sin estudios puedan tener la misma estructura de salarios, vista ésta como los ingresos promedio entre los diversos tipos de empleo, en

relación a aquellos con estudios medios o superiores. Simplemente porque los primeros pueden no tener “acceso” a la gama completa de tipos de empleo disponibles, mientras que los segundos sí.

En el caso de la interacción con el factor situación sindical, se comprueba un hecho que se observó previamente cuando se estudió la variabilidad del ingreso dentro de grupos homogéneos de personas por nivel educativo alcanzado. Para algunos niveles educativos se mantienen las no-diferencias en ingresos promedio, sugeridas en el modelo de análisis de varianza de efectos principales, entre aquellas personas sindicadas de las no sindicadas, mientras que en otros niveles parece haber diferencias significativas. Aquí el modelo de análisis de varianza completo sugiere realizar un estudio más profundo, por nivel educativo, de la situación específica de aquellas personas sindicadas de las que no lo están y por ende de la influencia del sindicato en los diferenciales de los ingresos. En la presente Tesis se había llevado a cabo un estudio por niveles educativos para investigar los factores causantes de las disparidades en los ingresos entre personas con el mismo grado escolar alcanzado, en la sección 3.4.3, ahí se observó que solamente para aquellos con educación secundaria y media superior el sindicato ejercía una influencia positiva en sus ingresos, mientras que para los otros grupos no. De esta forma, en el caso del mercado laboral del AMM se podría concluir que el sindicato no cumple con su función redistributiva, puesto que no beneficia a los grupos con menores grados educativos (los menos favorecidos), sino a aquellos con un nivel medio de educación.

Así, las interacciones analizadas estudian las diversas estructuras salariales que enfrentan los individuos pertenecientes a la PEA del AMM, en cada factor presentado, a través de los distintos niveles educativos. Su aportación al estudio de la desigualdad del ingreso radica en la búsqueda de cambios importantes en las estructuras salariales de un factor, en relación con otro, que nos permitan identificar diferencias en ingresos que el mercado laboral parezca no justificar. Como en el caso de los grupos con menores grados académicos, quienes no parecían ser apoyados por los sindicatos, mientras que aquellos con educación media sí recibían beneficios. Este hecho merece la atención de los

³⁴ Se menciona que “los diferenciales en ingreso” existen, porque ya se habían observado tanto en el análisis de regresión como en el análisis de varianza de efectos principales.

encargados de las políticas públicas, en su tarea de combate a las desigualdades económicas.

2. *La actividad económica*, es un factor interesante en sí mismo, en la determinación de los diferenciales en ingresos dentro de la PEA del AMM. Se le relaciona con la importancia de la elección de algún empleo en un área específica de la economía urbana. Es decir, se investiga a través de éste si realmente importa el lugar donde una persona decida trabajar (en este caso actividad económica), o no.

Dado que el efecto principal del factor sigue siendo significativo se puede afirmar que sí importa la actividad económica en la que uno decida trabajar, aunque no tan decididamente como el grado de estudio alcanzado por la persona, el cual vuelve a explicar una mayor proporción de la variabilidad del ingreso.

El factor no parece interactuar con: el estado civil, la situación sindical, el sexo ni con los grupos de edad. Por tanto, la estructura de las diferencias en ingresos promedio entre los distintos estados civiles se mantiene a través de todas las actividades económicas. Asimismo, en el caso de la situación sindical, parecen mantenerse las no-diferencias en ingresos entre aquellos sindicados de los que no los están, a lo largo de todas las actividades económicas del AMM, confirmándose el resultado global de no-diferencias en ingresos entre los sindicados y los no-sindicados, pero en esta ocasión a través de cada una de las actividades económicas analizadas. Por su parte, el factor sexo tampoco parece cambiar su relación de no-diferencias en ingresos, examinándose las posibles diferencias por género entre los sectores económicos del AMM, es decir, no se observan diferencias significativas en los ingresos promedio entre hombres y mujeres en ninguna de las actividades económicas analizadas. Tampoco parecen cambiar las diferencias en ingresos entre los grupos de edad a lo largo de las distintas actividades económicas.

Por otro lado, el factor actividad económica interactúa fuertemente con el tipo de empleo y la antigüedad en el último empleo. Lo cual significa que las diferencias en ingresos promedio entre los distintos tipos de empleo se modifican al movernos de una actividad económica a otra. Igual sucede dentro de los distintos niveles de antigüedad, las

diferencias en ingresos entre los grupos de antigüedad en algunas actividades pueden aumentar y en otras disminuir, es decir, en algunos sectores económicos se compensa más los años de antigüedad en el empleo, en relación a otros en donde los diferenciales son menores. No obstante, dado que el objetivo esencial del presente estudio es investigar si se presentan, o no, cambios en la estructura de los salarios de un factor a través de los niveles de los otros, enfatizando la utilidad de la metodología empleada, el análisis de descomposición de algunas interacciones, como la anterior, escapa a los límites del estudio.³⁵ Simplemente, es interesante constatar que en el caso del tipo de empleo y los rangos de antigüedad es importante elegir bien el sector de la actividad económica al cual se desea pertenecer, cuando la elección está basada precisamente en el tipo de empleo que se busca o las expectativas de ingresos futuros, creadas éstas principalmente por medio de los perfiles de ingreso según la antigüedad en el empleo.

Nótese que no se repite la interacción con la variable educación, puesto que conceptualmente es idéntica a la presentada con antelación. Así se procederá con el resto de los factores.³⁶

3. *El estado civil*; factor asociado con mayores ingresos para las personas casadas, en relación a las solteras, por cuestiones del mercado laboral.³⁷ Aquí vuelven a ser significativas las diferencias en ingresos entre los distintos estados civiles analizados, encontradas con antelación en el modelo de efectos principales. A su vez, parece no interactuar con los factores: grupos de edad, situación sindical y sexo. Es decir, no hay cambios significativos en los diferenciales en ingresos encontrados entre los grupos de edad al cambiar de estado civil. Tampoco se encuentran modificaciones en la relación de no-diferencias en ingresos promedio entre los sindicados y los no-sindicados, al examinarse los distintos estados civiles, situación similar entre hombres y mujeres al realizarse el mismo examen. Con lo cual, se confirma la relación de no-diferencias en ingresos dentro de los factores situación sindical y sexo, ya antes validada.

³⁵ El lector interesado en el comportamiento de los ingresos promedio de los diversos tipos de empleo, a través de las ramas de actividad económica, puede consultar las gráficas del apéndice 1.

³⁶ Véase la nota 30 para una ejemplificación.

³⁷ Véase Kenny (1983).

Los factores que sí interactúan con el estado civil son: el tipo de empleo y la antigüedad en el último empleo. En este caso las interacciones nos dicen que la estructura de salarios entre los diversos tipos de empleo varía al cambiarse de estado civil, al igual que los diferenciales en ingresos promedio entre los grupos de antigüedad, éstos no se mantienen iguales al pasar de un estado civil a otro.

Por tanto, se puede concluir que el estado civil juega un papel muy importante al momento de establecerse las diferencias en ingresos entre los diversos tipos de empleo y los rangos de antigüedad en un empleo.

4. El tipo de empleo; un factor que sin duda nos ayuda en la explicación de los diferenciales en ingresos entre las personas pertenecientes a la PEA del AMM. Su aportación directa es significativa y parece no interactuar con los cuatro factores restantes: la antigüedad, los grupos de edad, la situación sindical ni con el sexo. De esta forma, parecen no variar las diferencias en ingresos entre los diversos niveles de cada uno de estos factores, al examinarse a través de los distintos tipos de empleo. En el caso de la antigüedad en el último empleo y de los grupos de edad, parecen no variar las diferencias en ingresos entre las clases de cada factor cuando se examinan dentro de cada uno de los tipos de empleo. A su vez, en el caso de la situación sindical y el sexo, las no-diferencias en los ingresos dentro de cada uno de esos factores parecen mantenerse entre los distintos tipos de empleo, reforzando los resultados previos.

5. La antigüedad en el último empleo; no hay duda que la antigüedad en el empleo influye positivamente en los ingresos percibidos, algunas veces de forma más clara que en otras. En este caso su contribución directa sigue siendo significativa y analizando las interacciones con los últimos tres factores parece no haber algún efecto significativo conjunto con: los grupos de edad, el sexo ni con la situación sindical.

En el primer caso, la no-interacción nos dice que los diferenciales en ingresos entre los grupos de edad se mantienen a través de todos los rangos de antigüedad presentados, mientras que en el caso del sexo y la situación sindical el resultado es parecido a los previamente obtenidos, las no-diferencias en los ingresos promedio entre los niveles de

cada uno de estos factores se mantienen a lo largo de los diversos rangos de antigüedad. Por tanto, se confirman resultados preliminares.

6. *El sexo*; uno de los factores al cual se le ha dedicado un énfasis especial a lo largo del estudio. Los resultados obtenidos parecen no coincidir con lo comúnmente encontrado. Es decir, en ninguno de los casos analizados parecen existir diferencias en los ingresos por hora entre los hombres y mujeres pertenecientes a la PEA del AMM. Tanto en el modelo de efectos principales como en los análisis de varianza dentro de grupos homogéneos de población clasificados por su grado educativo. Este último modelo con interacciones no es la excepción, nuevamente no se encuentran diferencias en los ingresos promedio de ambos géneros.

Tampoco se observan interacciones importantes con los dos últimos factores por analizar: la situación sindical ni los grupos de edad. Así como no se encontraron interacciones significativas con los factores previamente estudiados. Por lo que se puede concluir que en ninguno de los casos presentados cambia la relación de no-diferencias en los ingresos promedio entre hombres y mujeres. Específicamente, en los dos últimos casos presentados, parece no importar si se está o no sindicado o el grupo de edad al que se pertenezca, las no-diferencias significativas en los ingresos se mantienen. Un hecho no aceptado del todo en la literatura económica, ni en la vida diaria, pero que en el presente estudio no existe duda alguna de su veracidad. Obviamente se deben recordar los comentarios sobre la muestra utilizada en el estudio, la cual representa a la PEA del AMM, un área económicamente fuerte. Además se refleja la situación vivida sólo en 1996. Por tanto, no se puede generalizar el resultado a nivel nacional, aunque se puede pensar en relaciones muy parecidas en otras áreas urbanas desarrolladas, como la Ciudad de México y Guadalajara.³⁸

7. *La situación sindical*; como se ha visto, este factor no parece influir de manera directa en los ingresos individuales de la PEA del AMM, en este último modelo parece ser que tampoco lo hace. Además, no interactúa con el último factor a estudiar: los

grupos de edad. Por lo tanto, parece no importar el grupo de edad al que se pertenezca, las no-diferencias entre los que están y no sindicados se mantienen.

Cabría hacer mención a la única interacción que fue válida para este factor, aquella con el factor educación, analizada en el primer apartado de esta sección. El resultado es interesante, parece que los sindicatos no están cumpliendo con su papel redistributivo del ingreso para apoyar a los que menos tienen. En realidad, se observa una mejora en los ingresos de aquellos con educación media, pero no se aprecian efectos positivos para aquellos con menores grados escolares.

8. Los grupos de edad; su contribución a la explicación de los diferenciales en ingresos ha quedado implícita en los análisis anteriores. Sería conveniente resaltar el hecho de la no-interacción entre este factor y algún otro. Es decir, los diferenciales en ingresos entre los grupos de edad, se mantienen a través de los diversos niveles de cada uno de los demás factores.

Este tipo de investigación constituye un útil ejercicio de abstracción que nos ayuda a entender mejor las estructuras salariales de un factor, tomando en consideración a los otros de forma explícita. Se prueba estadísticamente si los diferenciales, o no-diferenciales, entre diversas categorías clasificativas se mantienen en todo momento.

Cerramos esta parte de la investigación sobre las desigualdades del ingreso en Monterrey con un compendio de los resultados más interesantes. Entre los cuales destaca la trascendencia de la educación como principal herramienta en la búsqueda de una distribución del ingreso más equitativa. Vuelve a ser el factor con mayor peso explicativo en el modelo completo de análisis de varianza con interacciones. A su vez, no se detectaron interacciones significativas entre este factor y los factores actividad económica, estado civil, antigüedad en el último empleo, el sexo ni con los grupos de edad. Por tanto, se concluye que la importancia del capital humano a lo largo de todas las actividades económicas, los estados civiles, los rangos de antigüedad y de edad y entre hombres y mujeres es reconocida y premiada de manera muy similar dentro de las ramas

³⁸ Para el lector interesado, se presentan algunos análisis de varianza para las áreas urbanas del D.F. y

clasificadoras de cada uno de estos factores. De igual forma, se encontraron interacciones significativas entre la educación y los factores tipo de empleo y situación sindical, esperado en el primer caso por estar muy ligados la gama de tipos de empleos disponibles con el grado educativo, pero, no así en el segundo, dado que no se habían detectado diferencias entre los ingresos promedio de las personas sindicadas de las no sindicadas. Esperábamos que se mantuviera esta relación de no-diferencias dentro del factor situación sindical a través de los grados educativos, como de hecho sucedió entre las clasificaciones internas de los demás factores (demostrándose por medio de las interacciones correspondientes entre el factor situación sindical y cada uno de los otros factores), sin embargo, se encontró que sólo se mantenía para los grupos con educación superior y para aquellos con educación primaria, beneficiándose de la membresía del sindicato solamente las personas con educación secundaria y media superior. Este es un resultado conjunto del modelo de análisis de varianza completo con interacciones y un análisis de varianza para cada uno de los cuatro grupos educacionales.

Algo similar se observa en el caso del factor actividad económica, el cual no interactúa con los grupos de edad, el estado civil, la situación sindical ni con el sexo, indicando que la estructura de ingresos de estos factores (diferencias en los primeros dos casos, no-diferencias en los últimos) se mantiene constante a través de las ramas de la actividad económica. Mientras que en los casos del tipo de empleo y la antigüedad, estos sí interactúan con la actividad económica, por tanto, al cambiar de rama de la actividad económica cambian las diferencias internas de los factores. Punto de interés para aquellos que buscan empleo en algún área de la economía, basando su elección en los ingresos esperados por el tipo de empleo o el perfil de ingresos según la antigüedad en el empleo.

Finalmente tenemos el factor sexo. A través del análisis de las interacciones de éste con cada uno de los factores restantes concluimos que no se presentan diferencias en los ingresos promedio de los hombres y las mujeres pertenecientes a la PEA del AMM. Comparándose los ingresos promedio por género dentro de cada uno de los otros factores analizados, no se observaron diferencias significativas entre éstos, al no interactuar el sexo con algún otro factor, por tanto se confirman los resultados previos

del análisis de varianza de efectos principales y los estudios dentro de grupos homogéneos de la población clasificados por el nivel educativo alcanzado.

De esta forma, se demuestra la utilidad del análisis de varianza en aplicaciones económicas, principalmente en el estudio de la desigualdad. Hemos encontrado para el caso concreto del Área Metropolitana de Monterrey tres resultados importantes: la educación es el factor con mayor poder explicativo en las desigualdades del ingreso; no se identifican diferencias en los ingresos promedio entre hombres y mujeres; ni tampoco se presentan diferencias entre las personas sindicadas de las no-sindicadas a través de los demás factores, excepto entre las personas con educación media (secundaria y preparatoria) quienes sí se benefician del apoyo sindical.

Capítulo V. Conclusiones y Análisis Final

El estudio de la desigualdad del ingreso ha pasado por varias etapas evolutivas a través de las cuales se han empleado diversas técnicas para medir la desigualdad misma, así como sus determinantes. Una de las aportaciones principales de este estudio fue implementar una metodología de análisis de la desigualdad, medida ésta a través de sus factores componentes, muy utilizada en las ciencias sociales, pero, poco explotada en la economía: el Análisis de Varianza. Una herramienta estadística confiable y precisa, cuya aplicación en la investigación de los componentes de los diferenciales del ingreso nos ayudará a comprender mejor los mecanismos que interactúan en la determinación de dichas diferencias.

Uno de los objetivos principales del trabajo fue medir la contribución relativa de los distintos factores estudiados en la explicación de las divergencias de los ingresos individuales. Se observó, en todos los modelos analizados, que el último grado de estudio alcanzado es el factor con mayor peso explicativo entre la población económicamente activa del Área Metropolitana de Monterrey.

El resultado es interesante, pues confirma la importancia relativa del capital humano en una de las zonas económicas más importantes del país: Monterrey. Los cambios estructurales que ha vivido México le han dado al factor capital humano un papel protagónico como motor del crecimiento del país. La “globalización” de la economía y la cada vez mayor competitividad entre las industrias han propiciado que la demanda laboral se incline en favor de los más capacitados. Esa mayor demanda ha sobre-compensado el crecimiento de la oferta laboral con mejor educación y por ende los salarios han aumentado.³⁹

La importancia de la educación no es un asunto de orden secundario. Aunado a la creciente demanda por personal con mayores dotaciones de capital humano, nos encontramos con todos los problemas que enfrenta el sistema educativo nacional, los

³⁹ Véase Valero (1995) y Meléndez (1997a y 1997c).

cuales pueden conducir a un problema mayúsculo: una creciente desigualdad en el ingreso.⁴⁰

También se observa un grado de “madurez” del mercado laboral, en el sentido de no hacer distinciones en los ingresos por hora entre hombres y mujeres. Además, esta relación de no-diferencias se mantiene dentro de cada uno de los otros factores estudiados (lo cual se observa por medio del modelo con interacciones). Esto podría deberse a una autoselección de las mujeres pertenecientes a la PEA, las cuales pueden ser las más capaces para trabajar y por ende logran ingresos iguales (e incluso mejores) que los hombres en varias de las categorías explicativas analizadas.

Por otra parte el modelo de análisis de varianza con interacciones probó si existían diferencias en las estructuras de los ingresos de un factor (por ejemplo, el grado educativo alcanzado) dentro de cada uno de los niveles de los otros factores (por ejemplo, hombre o mujer en el factor sexo). Se observó una baja modificación en dichas estructuras, es decir, las diferencias entre los niveles de un factor se mantenían a lo largo de los niveles de los demás factores. Esto, dependiendo el factor estudiado, arroja información interesante para aquellos encargados de apoyar a ciertos grupos objetivo. Por ejemplo, dado que las no-diferencias entre hombres y mujeres se mantuvieron a través de todos los niveles de los otros factores, no se recomendarían acciones públicas encaminadas a reducir los posibles diferenciales en ingresos entre hombres y mujeres, puesto que no se observan.⁴¹ En cambio, parece ser que en el caso de los diferenciales por nivel educativo alcanzado, éstos cambian entre las personas sindicadas de las que no lo están. Resultado que se apoya en el análisis llevado a cabo en la sección 3.4.3, donde se observó que existían diferencias en los ingresos promedio entre aquellas personas que se encontraban sindicadas de las que no lo estaban, siendo las primeras las que mayores ingresos obtenían, dentro del grupo de personas con educación secundaria o media superior. Así, el sindicato no parece ser útil como medio redistributivo del ingreso, puesto que los menos educados no se benefician al estar sindicados, mientras que sí lo

⁴⁰ Véase Ornelas (1995) y Guevara (1992) para un estudio detallado del sistema educativo mexicano y sus problemas.

⁴¹ Por supuesto tomando en cuenta otros factores, como la existencia de posibles barreras desde el momento mismo de solicitar empleo.

hacen aquellos con educación media. Por lo tanto, el modelo de análisis de varianza completo nos proporciona la pauta a seguir en la investigación de los diferenciales en ingresos por factores componentes.

Al tener una sociedad con un poder adquisitivo elevado se fortalece la demanda interna de consumo y el ahorro. Generándose a su vez un círculo virtuoso en el sector productivo al incrementarse sus mercados potenciales y crearse adicionalmente mayores recursos para la inversión, vía mayor ahorro interno. Por ello, estudiando los determinantes básicos de estas disparidades se contribuye a la búsqueda de soluciones igualitarias que tiendan a fortalecer, no sólo la economía de los más necesitados, sino procurar el fortalecimiento de la economía nacional.

Por último, es importante señalar que los resultados aquí presentados se basan en una muestra obtenida de una área metropolitana. Sería conveniente ver el marco global incluyendo los diferenciales por regiones urbanas y rurales. Tomando en cuenta el factor migración.

De hecho esperaríamos encontrar que para las zonas urbanas la educación y la edad, o la experiencia, serían las variables importantes en la explicación de los diferenciales en el ingreso. Mientras que en las zonas rurales, el tipo de zona donde se vive (por ejemplo, la fertilidad del suelo) explicaría aún más los diferenciales en el ingreso.

Esto abrirá un prometedor campo de investigación sobre la desigualdad, particularmente en el sector rural entre los empleados, desarrollando procedimientos que tomen en cuenta la distribución de los dueños de los factores no laborales de la producción (tierra y capital) y ampliando el rango de estudio a escala nacional, entre regiones.

Bibliografía.

Adelman, Irma y Cynthia Taft Morris, (1973). "Economic Growth and Social Equity in Developing Countries," Standford University Press.

Camero Cortés, Cipactli, (1995). "Análisis de segregación y diferencias en el rendimiento a la escolaridad de acuerdo al sexo. Un estudio para el mercado laboral del Área Metropolitana de Monterrey," Tesis de Licenciatura, Facultad de Economía, Universidad Autónoma de Nuevo León.

Fei, J. C. H. y Kuo, S. W. Y., (1978). "Growth and the Family Distribution of Income by Factor Components," Quartely Journal of Economics, February, Vol 92, pp. 17-53.

Fields, Gary S. y Fei, J. C. H., (1978). "On Inequality Comparisons," Econometrica, March, pp. 303-316.

Fields, Gary S., (1979). "Decomposing LDC Inequality," Oxford Economic Papers, Vol.31, No.3, November, pp. 437-459.

_____, (1980a). "Poverty, Inequality, and Development," Cambridge University Press.

_____, y T. Paul Schultz, (1980b). "Regional Inequality and Others Sources of Income Variation," Economic Development and Cultural Change, Vol. 28, No 3, April, pp. 447-467.

_____, y T. Paul Schultz, (1982). "Income Generating Functions in a Low Income Country: Colombia," The Review of Income and Wealth, series 28, No.1, March, pp. 71-87.

Garro, N. y Llamas, I., (1995). "El rendimiento de la escolaridad, la capacitación y la experiencia laboral de los trabajadores del AMM," Ensayos, Facultad de Economía, Centro de Investigaciones Económicas, Universidad Autónoma de Nuevo León, Vol.XIV, No.1, mayo, pp. 79-106.

Gatica, Jaime, Alejandra Mizala, y Pilar Romaguera, (1995). "Interindustry Wage Differentials in Brazil," Economic Development and Cultural Change, Vol. 43, No. 2, January, pp. 315-331.

Gibbons, R. y L. Katz, (1992). "Does Unmeasured Ability Explain Inter-Industry Wage Differences?" Review of Economic Studies 59, July, pp. 515-535.

- Gollás, Manuel, (1982). "La Economía Desigualdad, Empleo y Distribución en México," Consejo Nacional de Ciencia y Tecnología, Circuito Cultural, Centro Cultural Universitario.
- Grilches, Zvi, (1977). "Estimating the Returns to Schooling: some Econometric Problems," *Econometrica*, Vol. 45, No.1, January, pp. 1-22.
- Groshen, E., (1991). "Source of Intra-Industry Wage Dispersion : How Much Do Employers Matter?" *Quartely Journal of Economics*, Vol. 106, pp. 869-884.
- Guevara Niebla, Gilberto, compilador, (1992). "La Catástrofe Silenciosa," Fondo de Cultura Económica. México.
- Juhn, Chinhui, y Kevin M. Murphy. (1997), "Wage Inequality and Family Labor Supply," *Journal of Labor Economics*, Vol.15, No.1, pp. 72-97.
- Katz, L., (1988). "Some Recent Developments in Labor Economics and Their Implications for Macroeconomics," *Journal of Money, Credit, and Banking*, 20, No. 3, August, pp. 507-530.
- Kenny, Lawrence W., (1983). "The Accumulation of Human Capital During Marriage by Males," *Economic Inquiry*, Vol.XXI, April, pp. 223-231.
- Kruger, A. y L. Summers, (1988). "Efficiency Wages and the Inter-Industry Wage Structure," *Econometrica*, Vol. 56, No. 2, March, pp. 259-293.
- Kuznets, Simon, (1955). "Economic Growth and Income Inequality". *American Economic Review*, March, pp. 1-28.
- Kuznets, Simon, (1963). "Quantitative Aspects of the Economic Growth of Nations. VIII. Distribution of Income by Size," *Economic Development and Cultural Change*, Vol. 11, No. 2, pt. 2, January, pp. 1-80.
- López Garza, Edgar, (1983). "Desigualdad, Capital Humano y Patrones de Discriminación en el Mercado Laboral. Un Estudio para el Área Metropolitana de Monterrey," Facultad de Economía, Centro de Investigaciones Económicas - Universidad Autónoma de Nuevo León.
- Levy, Santiago, (1994). "La Pobreza en México," en Félix Vélez (comp.), *La Pobreza en México: causas y políticas para combatirla*, México, Instituto Tecnológico Autónomo de México y El Trimestre Económico, Fondo de Cultura Económica, Serie Lecturas, No. 78, pp. 15-112.

Master, S. H., (1969). "Wage and Plant Size: An Interindustry Analysis," Review of Economics and Statistics, Vol. 51, pp. 341-345.

Meléndez Barrón, Jorge, (1995). "Evaluación de un Curriculum de Educación Universitaria: Estimaciones para el Área Metropolitana de Monterrey," Ensayos, Facultad de Economía, Centro de Investigaciones Económicas, Universidad Autónoma de Nuevo León, Vol.XIV, No.1, mayo, pp. 21-56.

_____, (1997a). "Inversión en Educación y Redistribución del Ingreso en México," Entorno Económico, Vol. XXXV, No. 211, noviembre-diciembre, Centro de Estudios Económicos, Universidad Autónoma de Nuevo León, pp. 9-12.

_____, (1997b). "La Rentabilidad de la Inversión en Escolaridad en el Área Metropolitana de Monterrey, 1993-1996," trabajo presentado en el Cuarto Encuentro Internacional de "Capital Humano, Crecimiento, Pobreza: Problemática Mexicana." Facultad de Economía, Universidad Autónoma de Nuevo León, 17 y 18 de septiembre.

_____, (1997c). "Demanda de Educación y Estructura Salarial: Evidencia del Efecto de la Habilidad para enfrentar el Desequilibrio en el Área Metropolitana de Monterrey," en Tijerina y Meléndez, editores. Tercer encuentro internacional "capital humano, crecimiento, pobreza: problemática mexicana," Universidad Autónoma de Nuevo León, pp. 201-232.

Mincer, Jacob, (1974). "Schooling, Experience, and Earnings," New York, National Bureau of Economic Research.

Mizala, Alejandra y Pilar Romaguera, (1998). "Wage Differentials and Occupational Wage Premia: Firm - level Evidence for Brazil and Chile," The Review of Income and Wealth , Series 44, No. 2, June, pp. 239-257.

Murphy, Kevin M., y Finis Welch, (1990). "Empirical Age-Earnings Profiles," Journal of Labor Economics, Vol.8, No.2, pp. 202-229.

Neter, John, Michael H. Kutner, Cristopher J. Nachtsheim, y William Wasserman, (1996). "Applied Linear Statistical Models, Fourth Edition," Richard D. Irwin, Inc., Times Mirror Higher Education Group, Inc.

Ornelas, Carlos, (1995). "El Sistema Educativo Mexicano," Fondo de Cultura Económica, México.

Rupert, Peter, Eric Severance-Lossin, y Erin Turner, (1996). "Earnings, Education, and Experience," Economic Review, 4th quarter, pp. 2-12.

SAS System for Linear Models, Third Edition, (1991) by SAS Institute Inc., Cary, NC, USA.

Searle, S. R., (1971). "Linear Models," New York, NY: Wiley.

Shorrocks, Anthony F., (1982). "Inequality Decomposition by Factor Components," Econometrica, Vol.30, No.1, January, pp. 193-212.

_____ (1983). "The Impact of Income Components on the Distribution of Family Incomes," The Quarterly Journal of Economics, Vol.98, No.2, May, pp. 311-326.

Schweitzer, Mark E., (1997). "Workforce Composition and Earnings Inequality," Economic Review, 2nd quarter, pp. 13-24.

Slichter, S., (1950). "Notes on the Structure of Wages," Review of Economics and Statistics, Vol. 32, pp. 80-91.

Todaro, Michael P., (1989). "Economic Development in the Third World, Fourth Edition," Longman Inc.

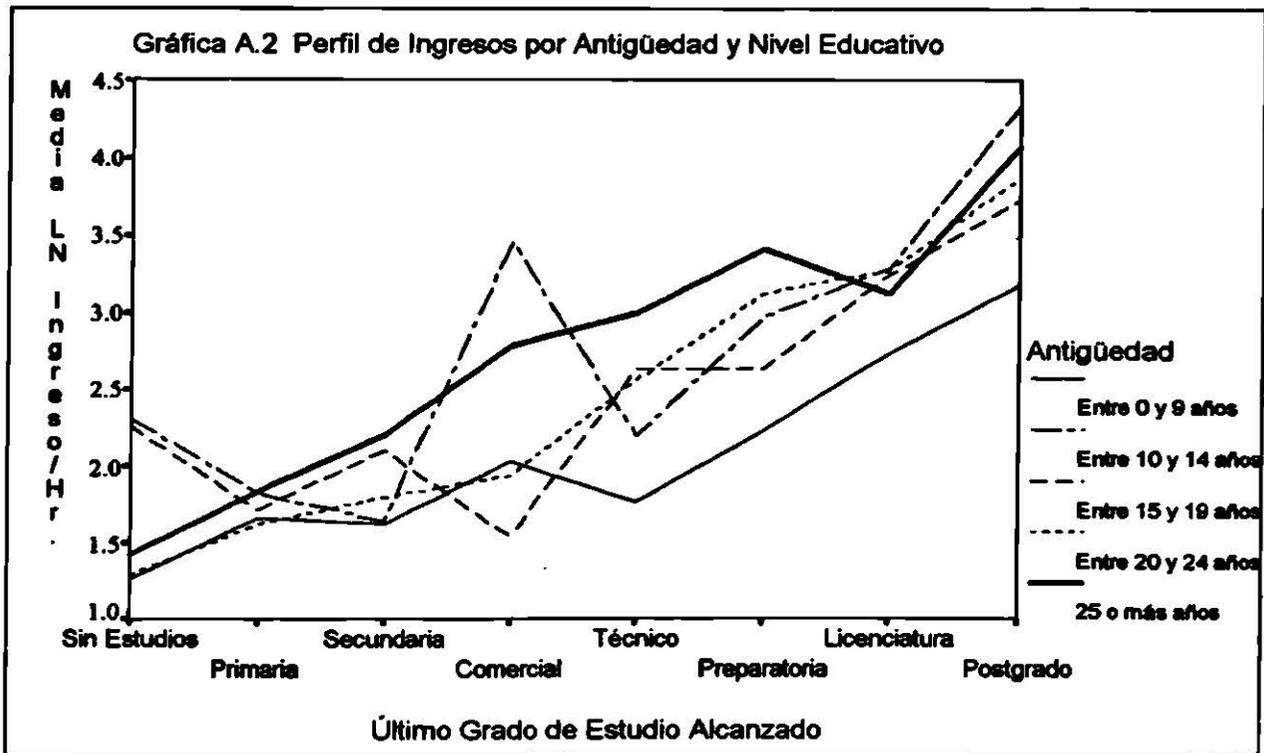
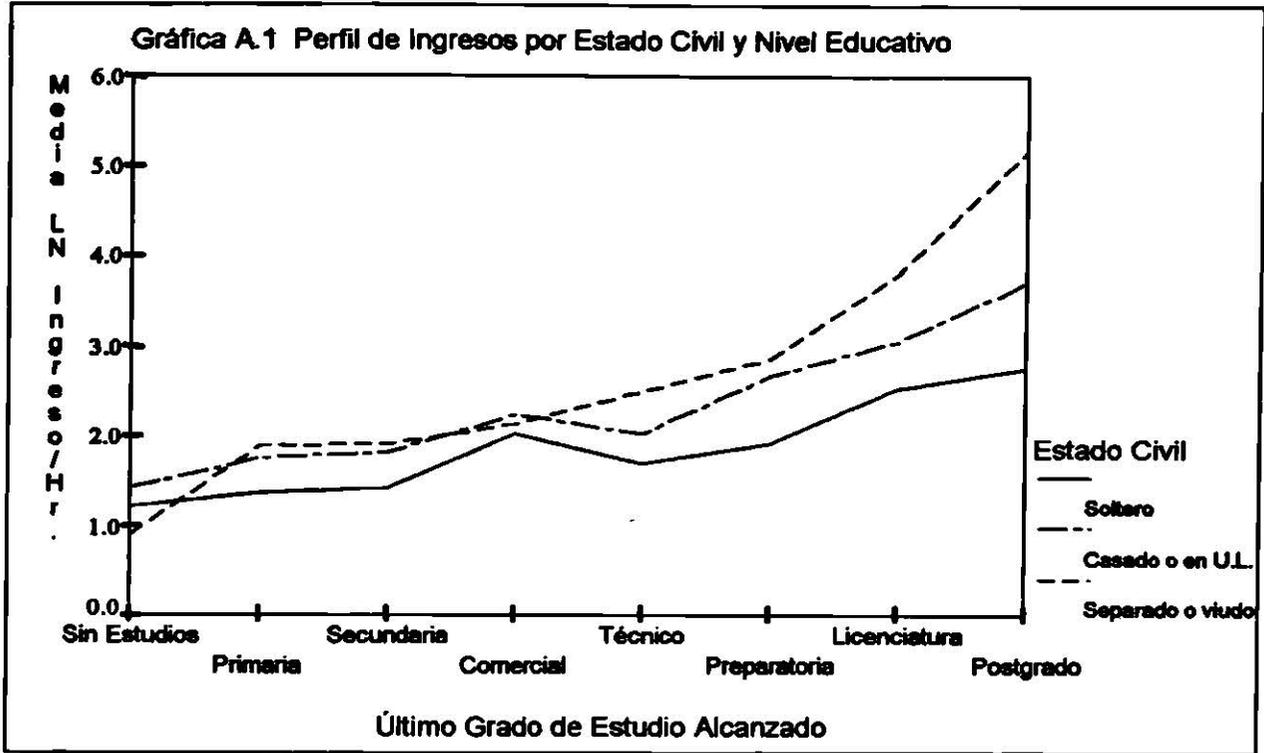
Valdez Moreno, Francisco, (1995). "Condiciones del Mercado Laboral para Hombres y Mujeres en el AMM. El Asunto de la Discriminación Salarial por Sexo," Tesis de Licenciatura, Facultad de Economía, Universidad Autónoma de Nuevo León.

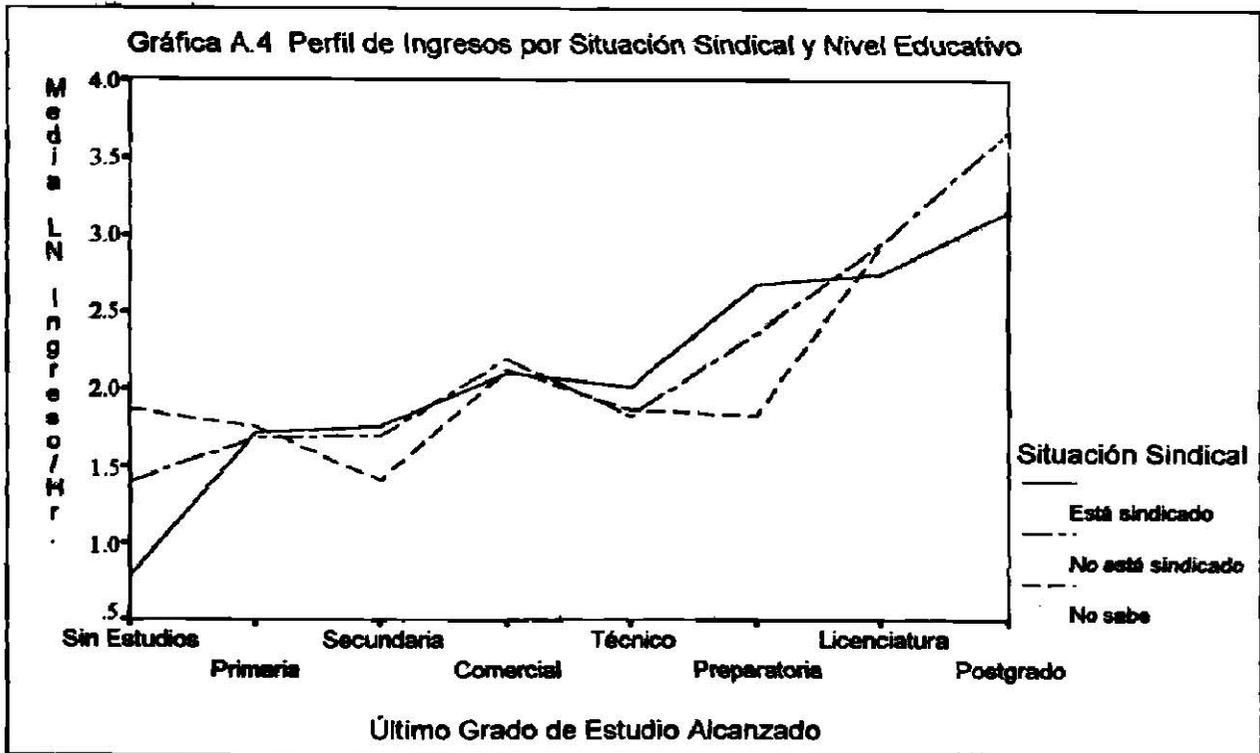
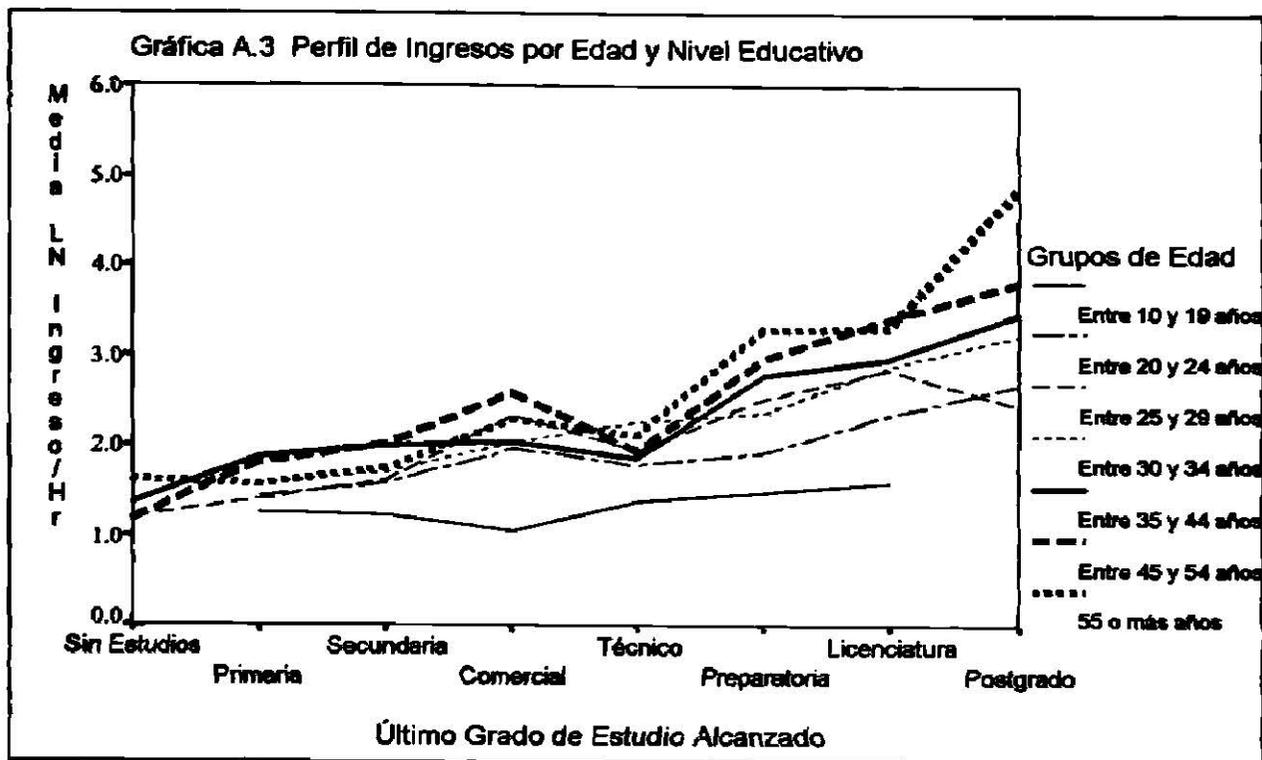
Valero Gil, Jorge N., (1995). "Notas para el Estudio de los Cambios en los Salarios Relativos 1979-1993 en el Área Metropolitana de Monterrey," Ensayos, Facultad de Economía, Centro de Investigaciones Económicas, Universidad Autónoma de Nuevo León, Vol.XIV, No.2, noviembre, pp. 41-67

Vellinga, Menno y Edgar López, (1986). "Ingreso: Distribución y Redistribución, Movilidad Social y Niveles de Vida, 1965-1985". Centro de Investigaciones Económicas. Universidad Autónoma de Nuevo León.

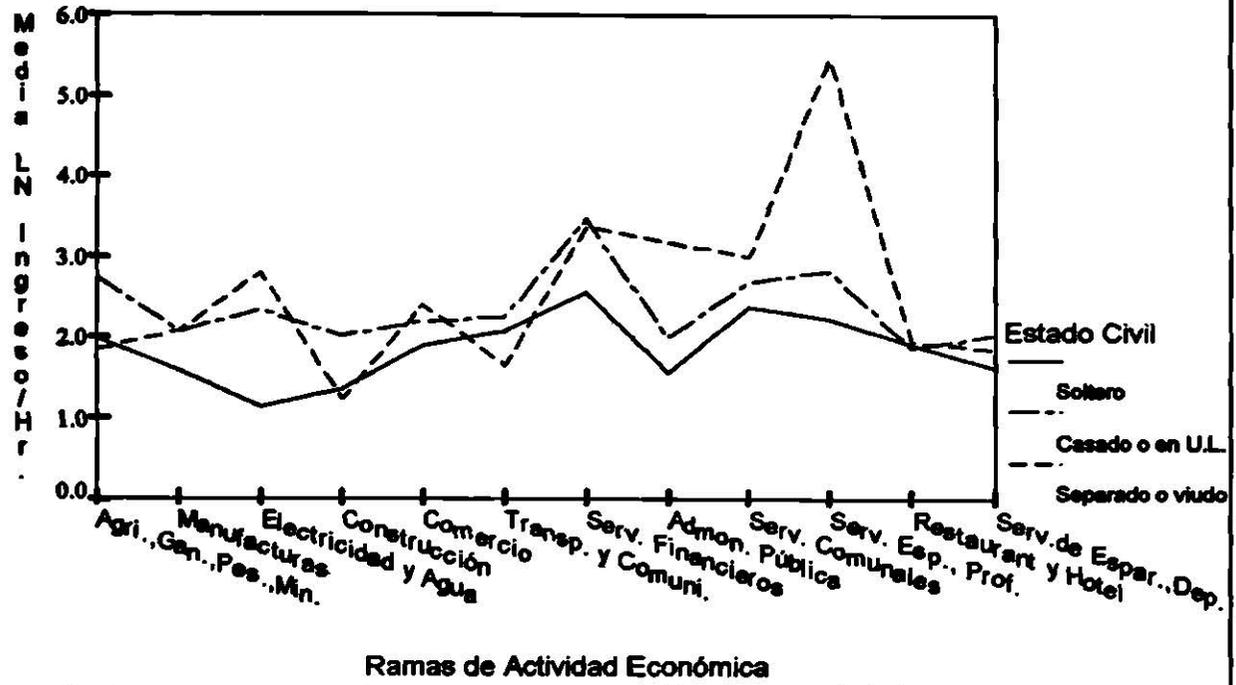
Williamson, Jeffrey, (1965). "Regional Inequality and the Process of National Development: A description of the Patterns," Economic Development and Cultural Change, Vol. 13, No. 4, Pt. 2. July, pp. 33-84.

Apéndice 1. Gráficas adicionales

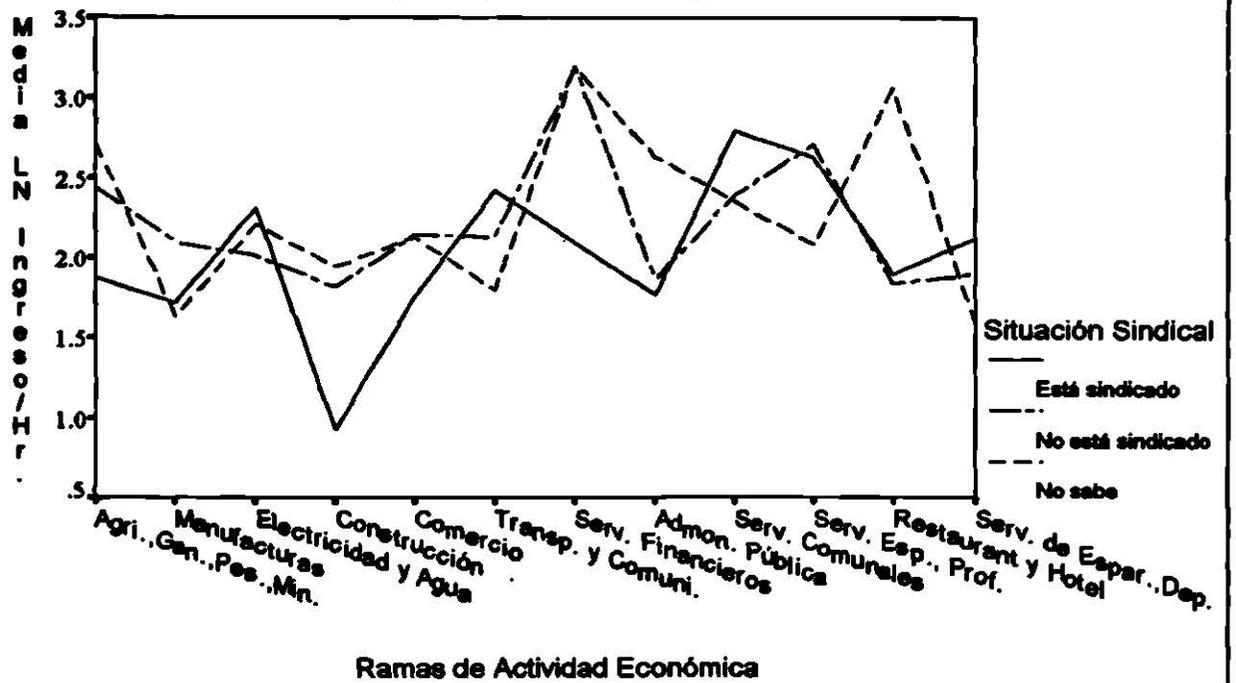




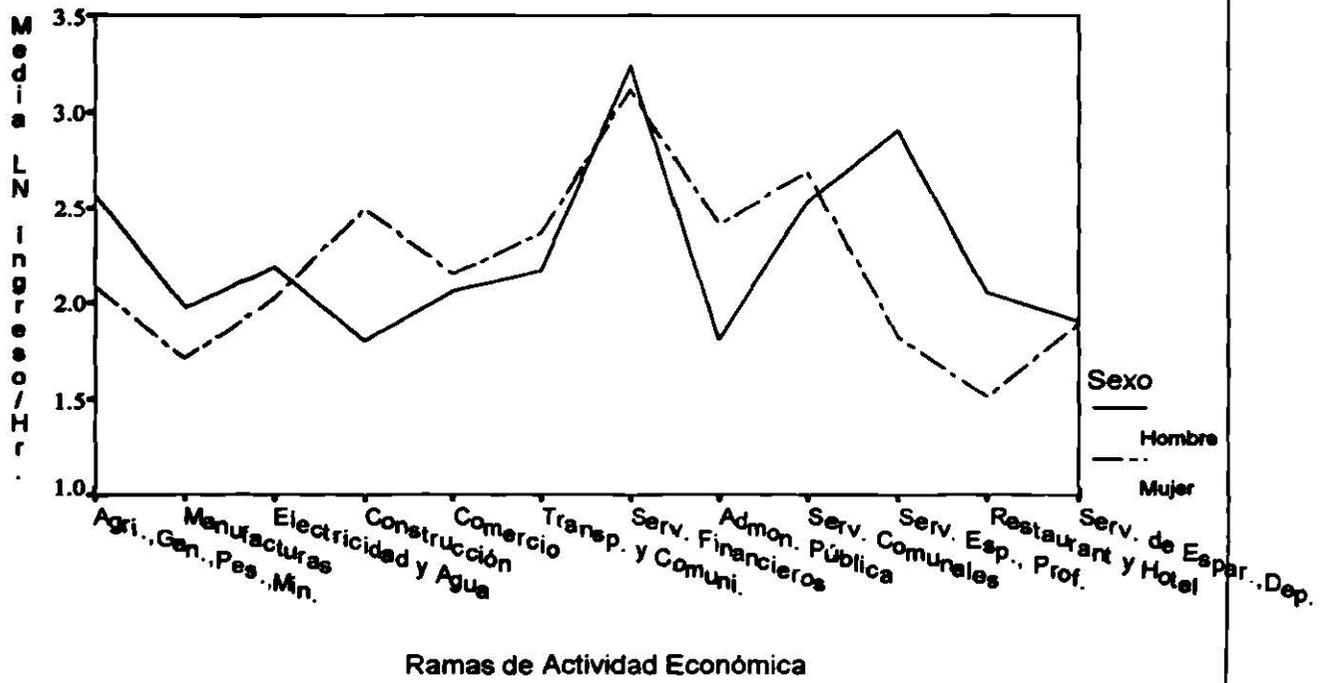
Gráfica A.7 Perfil de Ingresos por Estado Civil y Actividad Económica



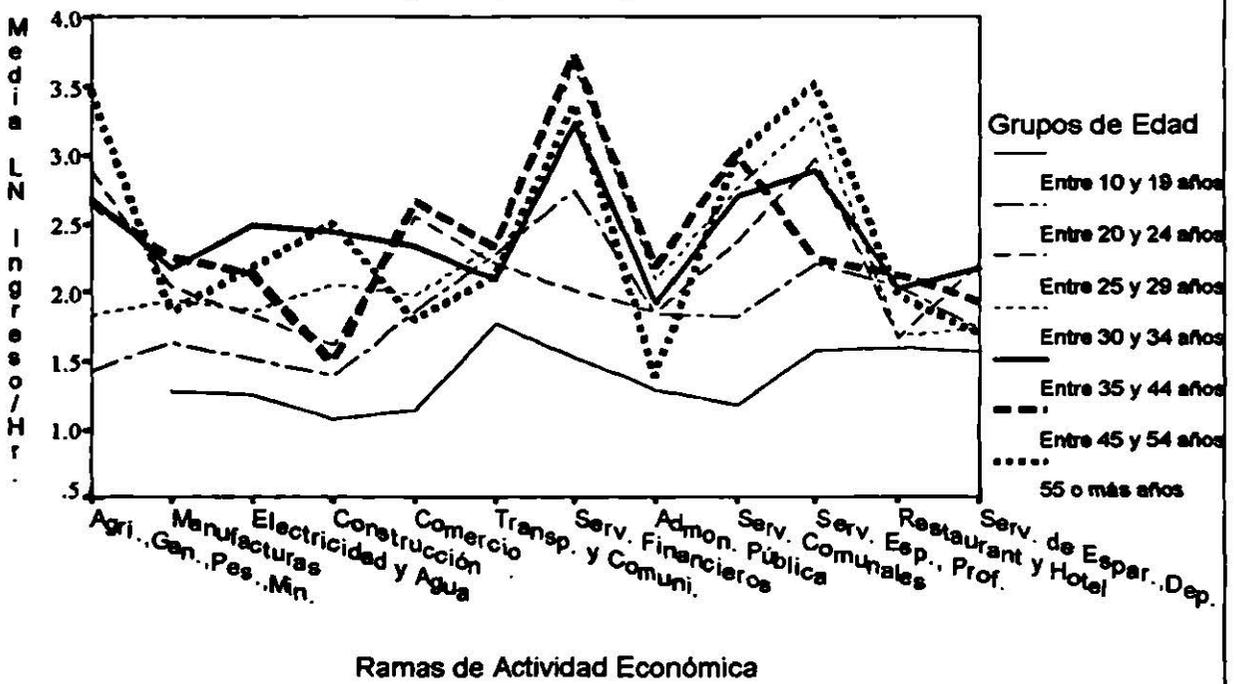
Gráfica A.8 Perfil de Ingresos por Sit. Sindical y Actividad Económica

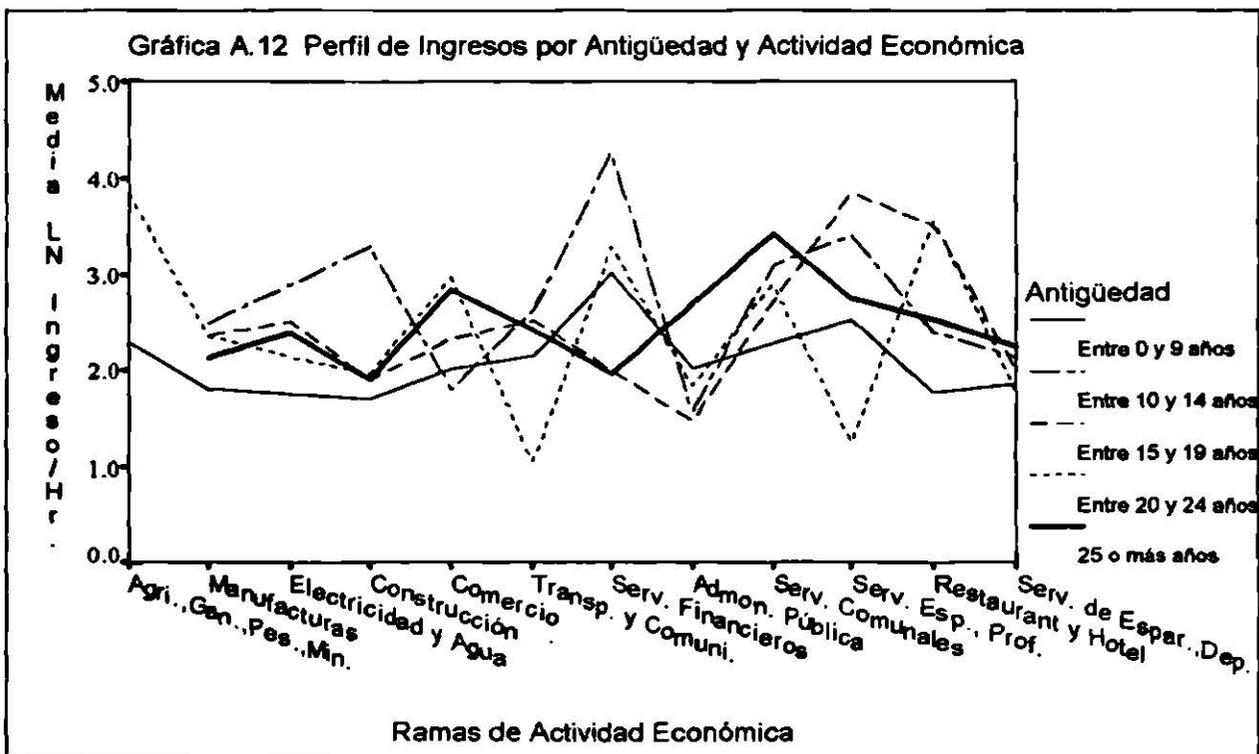
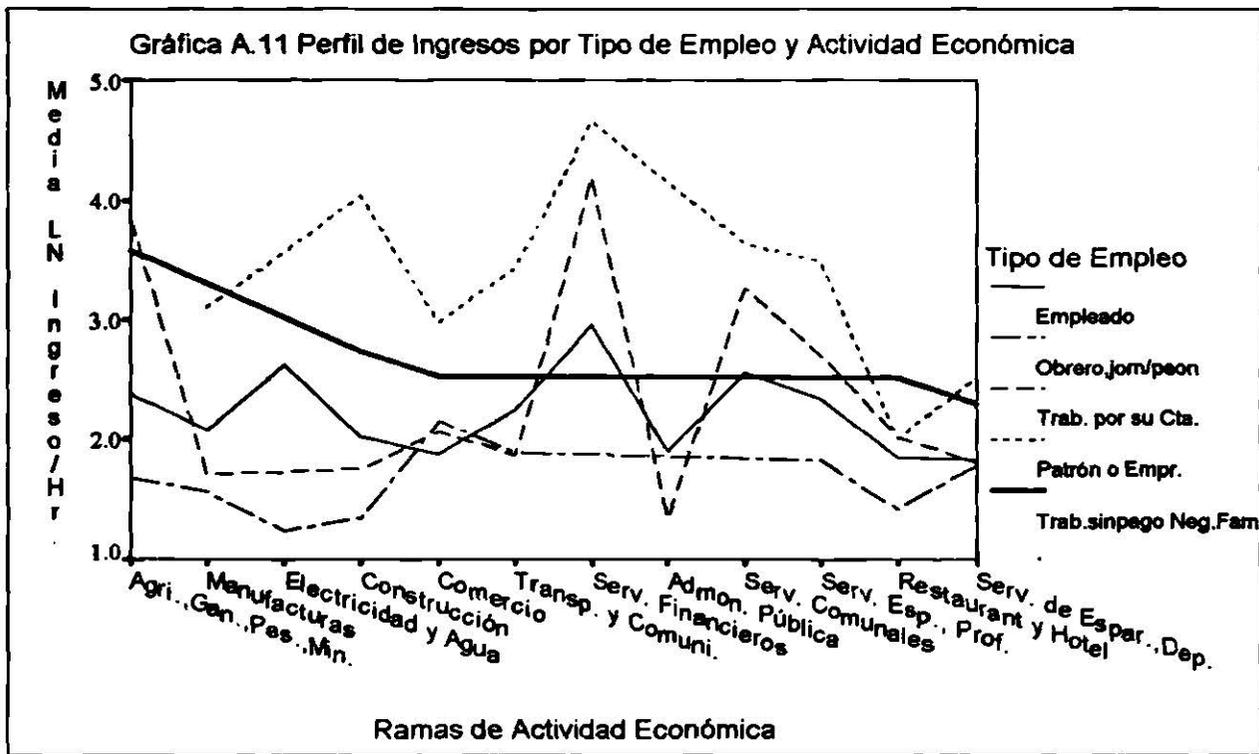


Gráfica A.9 Perfil de Ingresos por Sexo y Actividad Económica

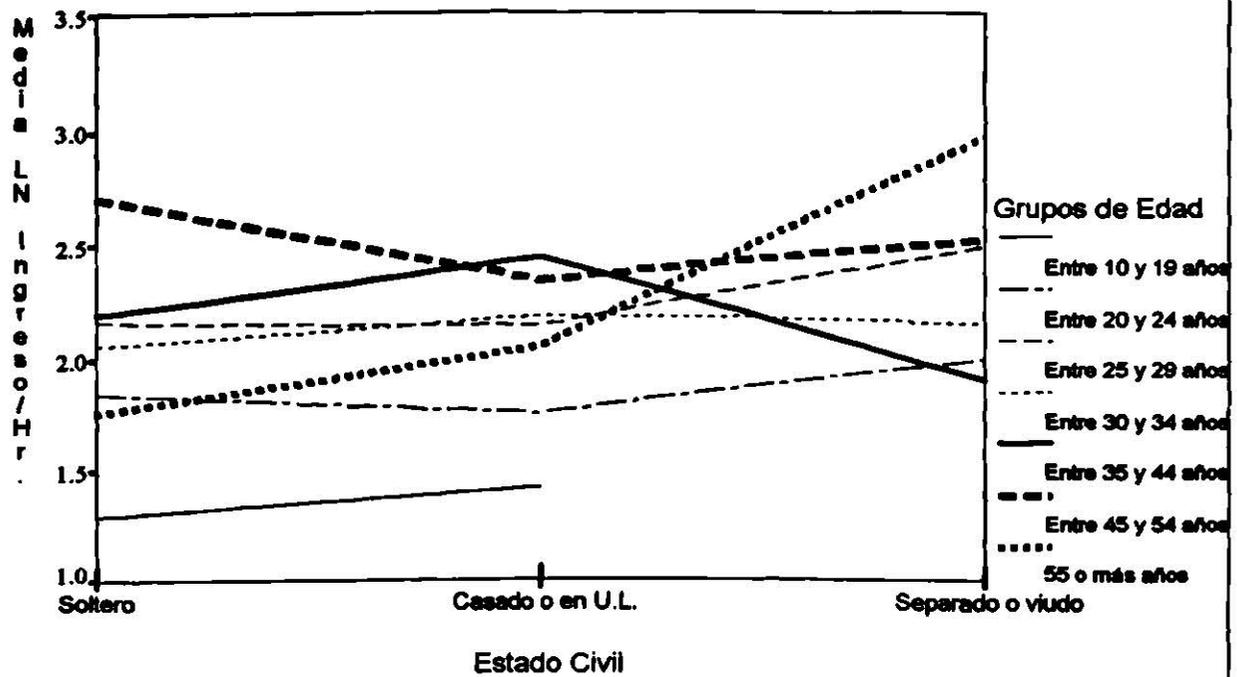


Gráfica A.10 Perfil de Ingresos por Edad y Actividad Económica

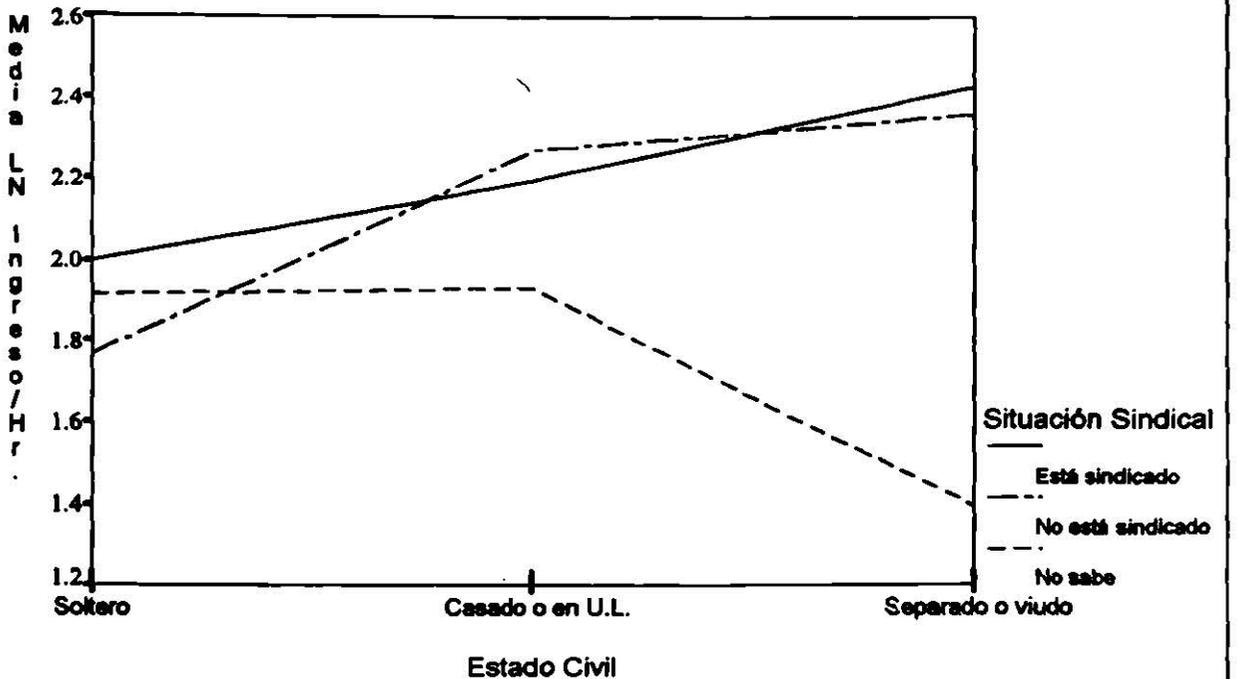




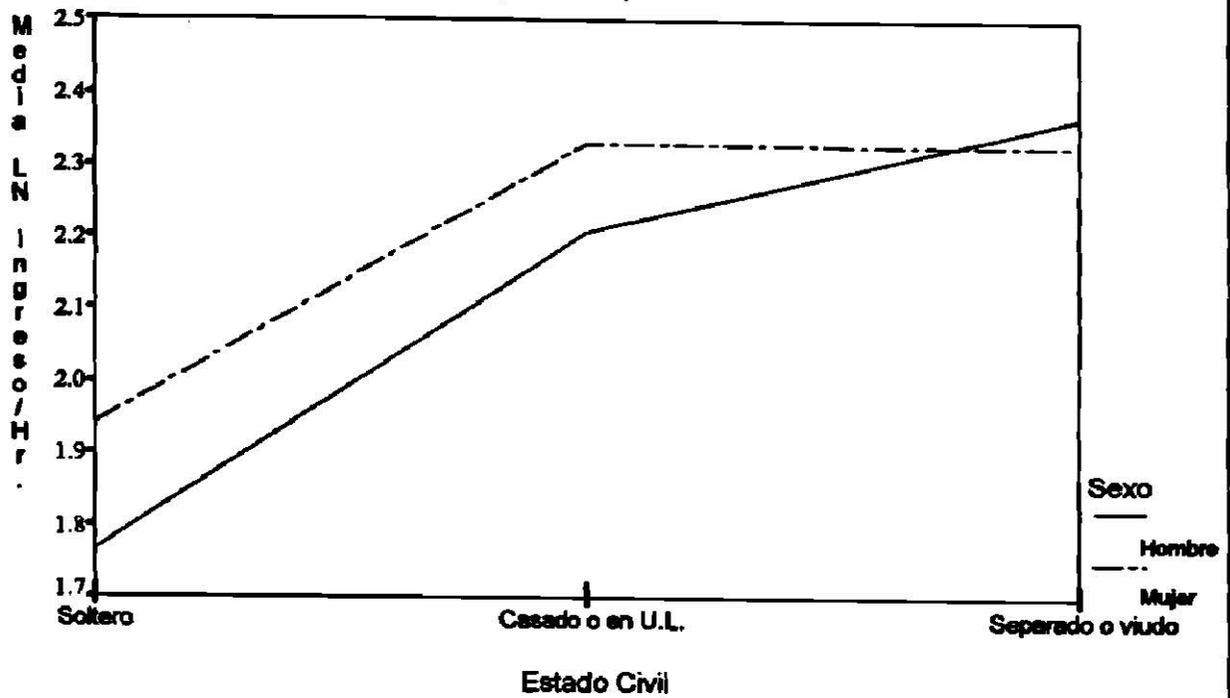
Gráfica A.13 Perfil de Ingresos por Edad y Estado Civil



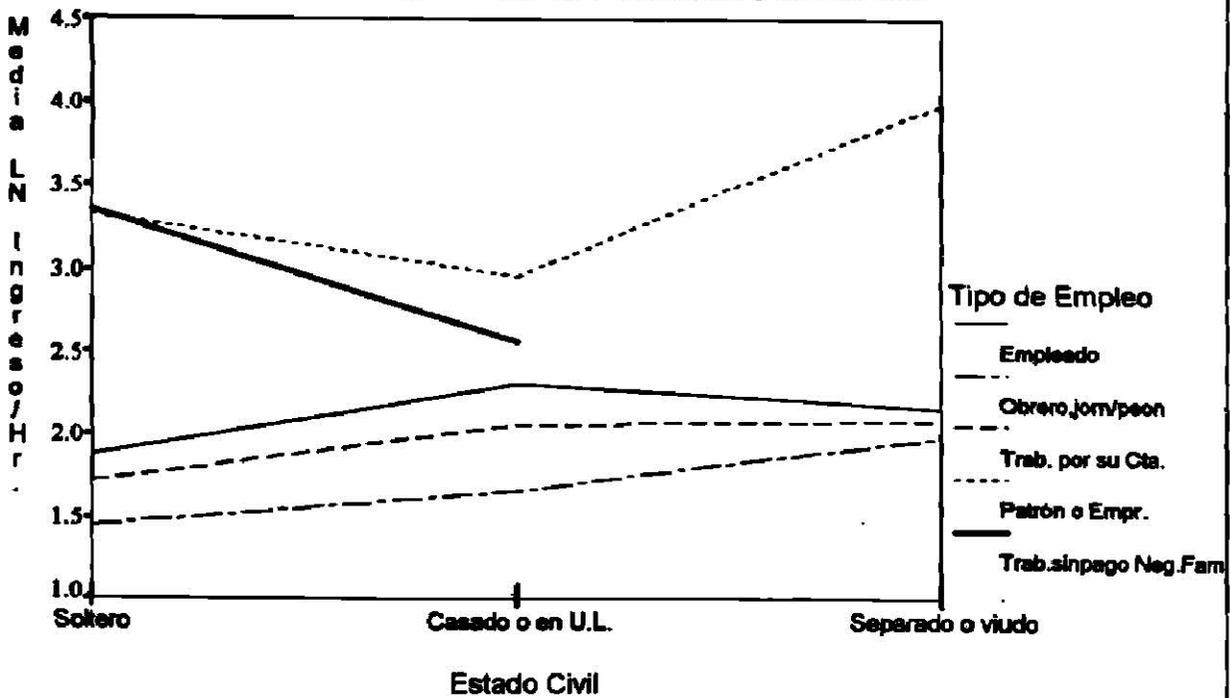
Gráfica A.14 Perfil de Ingresos por Situación Sindical y Estado Civil



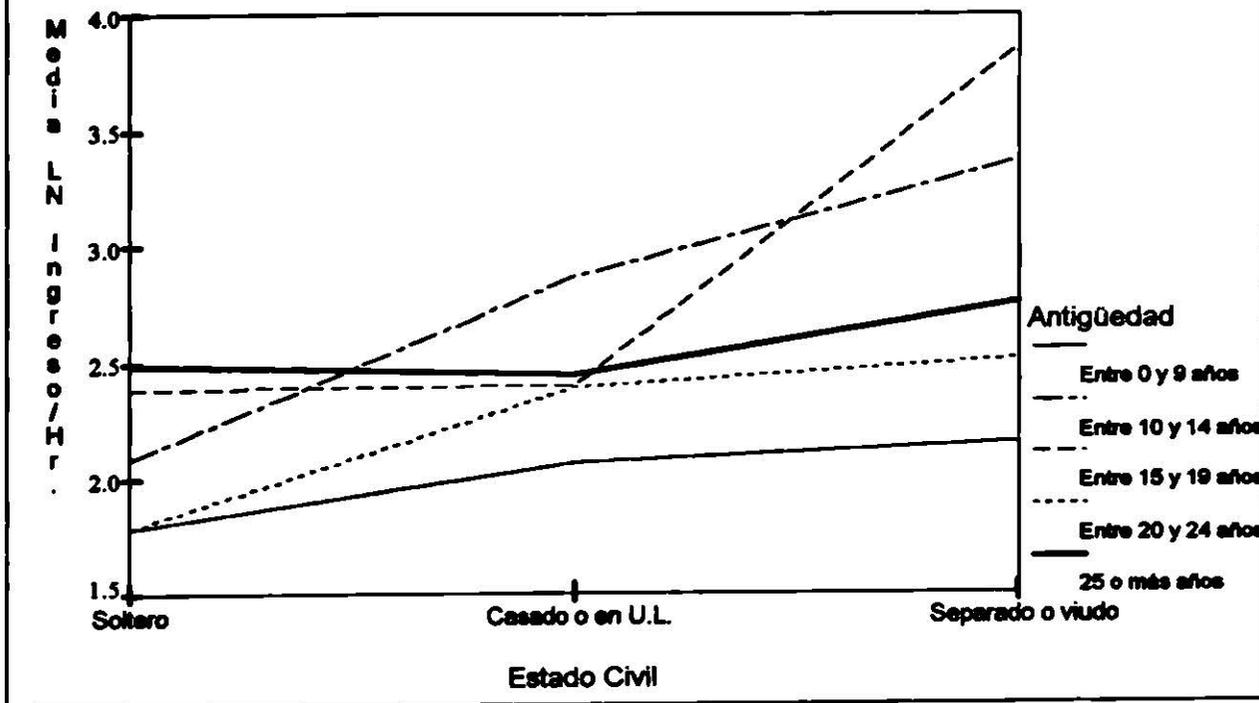
Gráfica A.15 Perfil de Ingresos por Sexo y Estado Civil



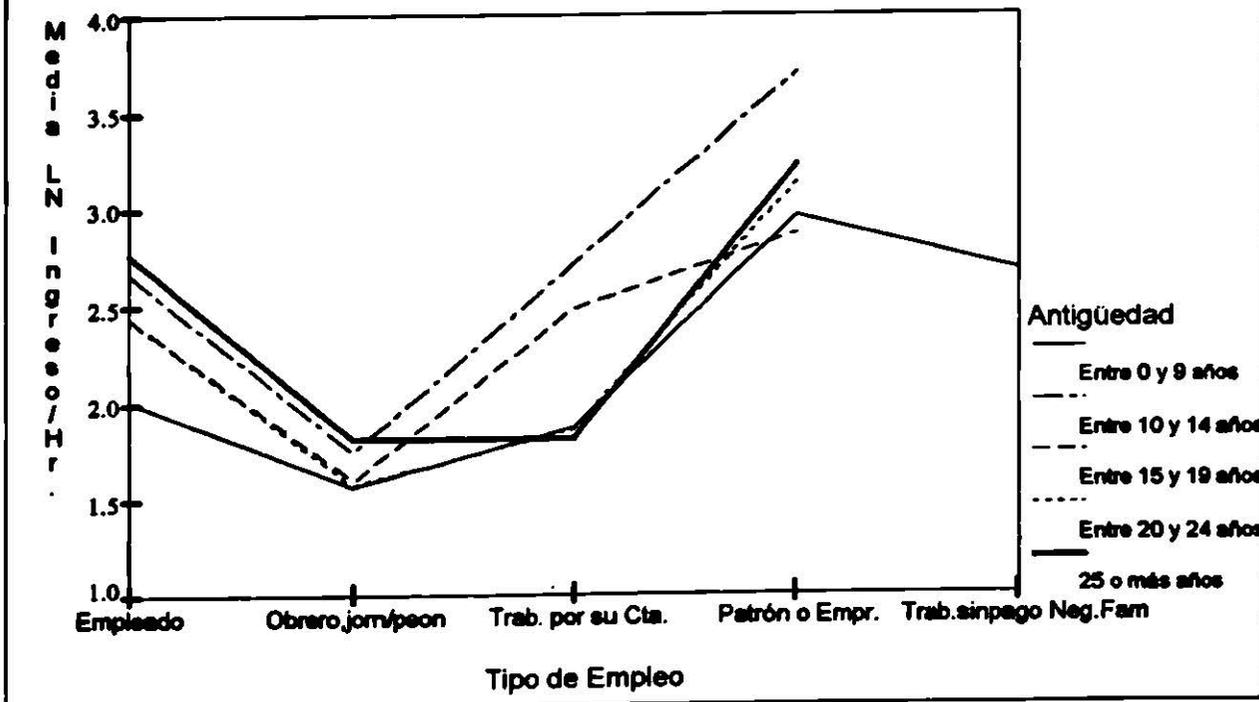
Gráfica A.16 Perfil de Ingresos por Tipo de Empleo y Estado Civil

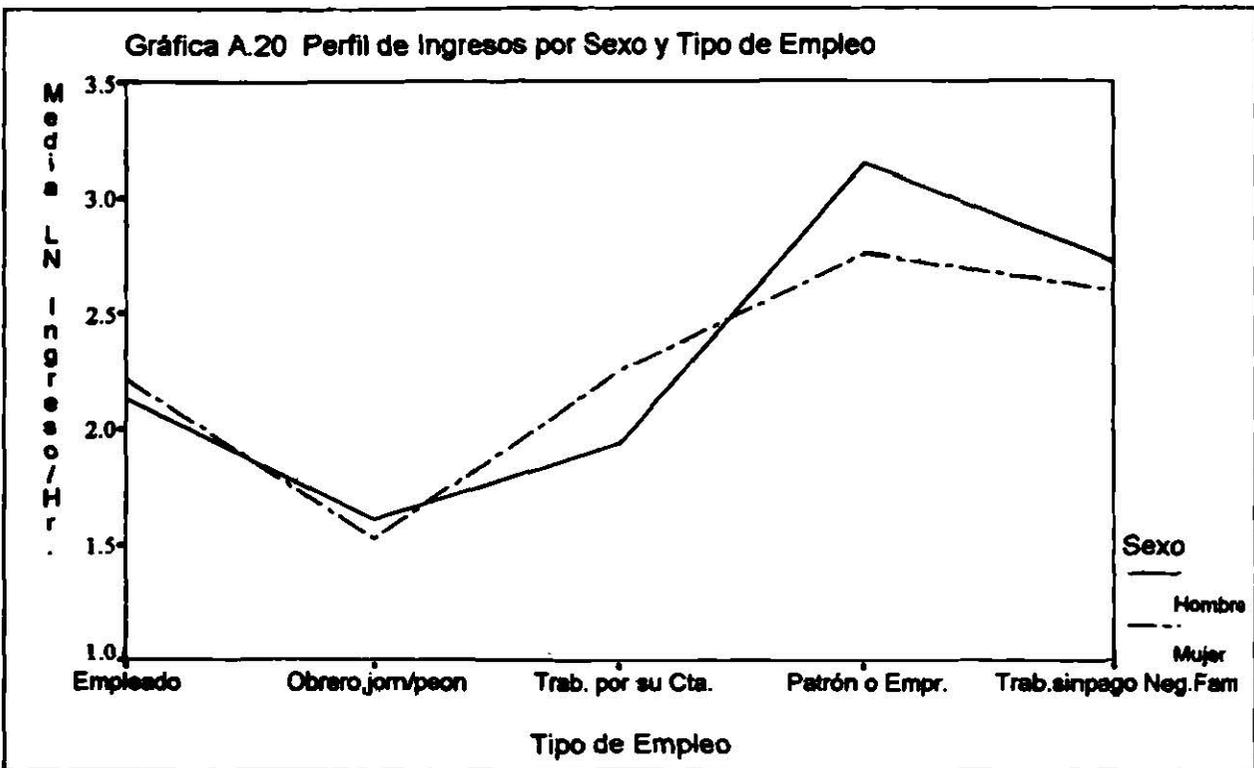
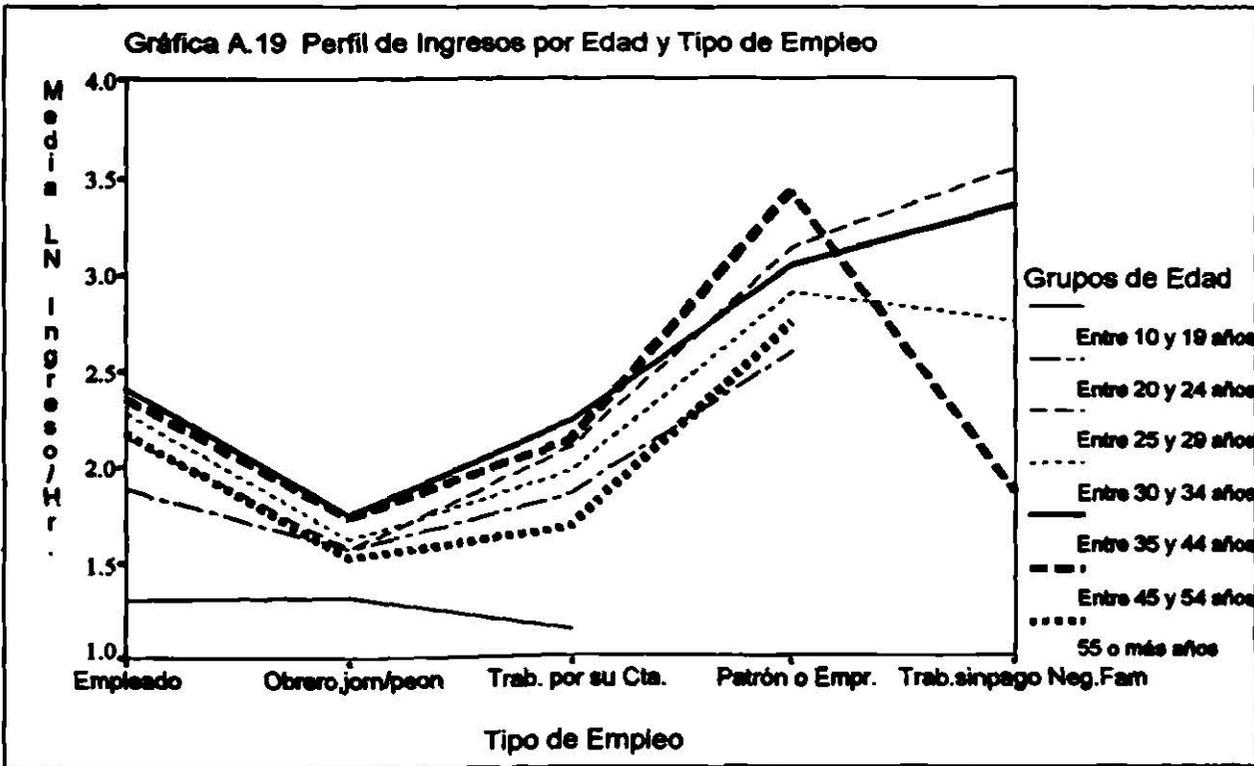


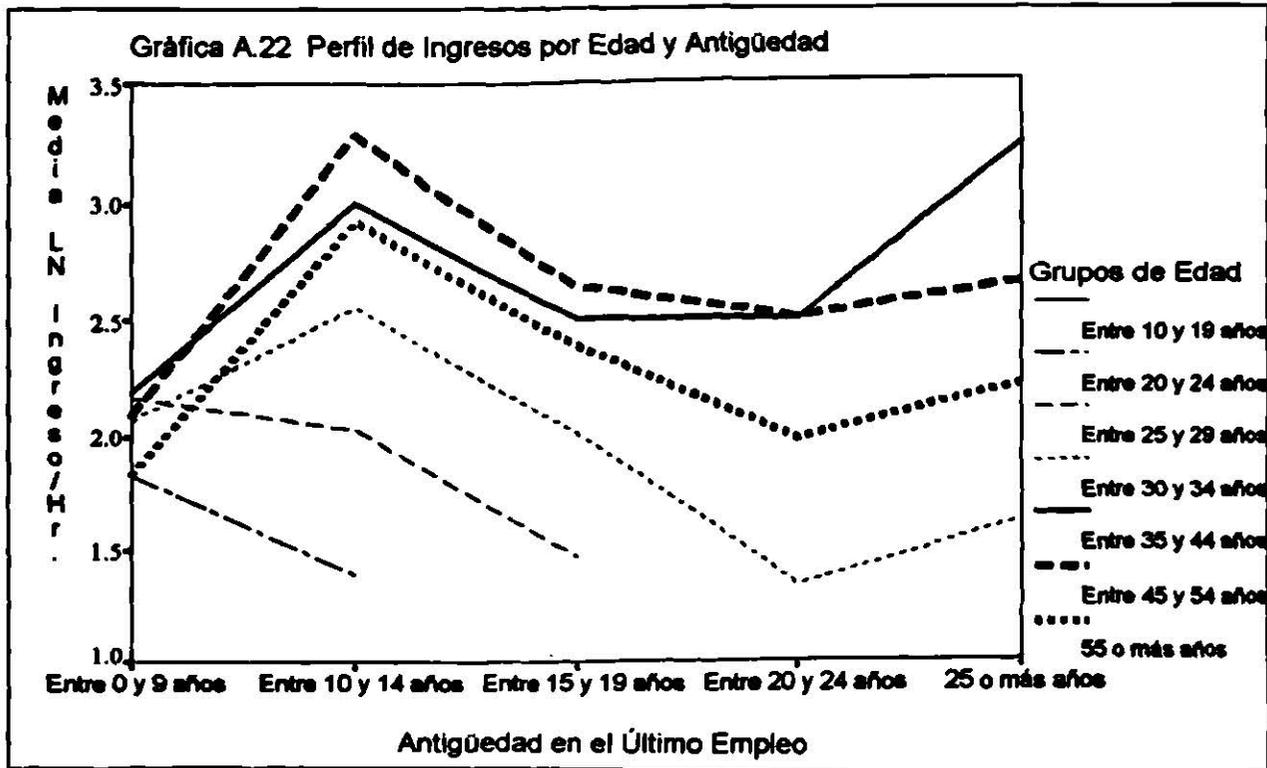
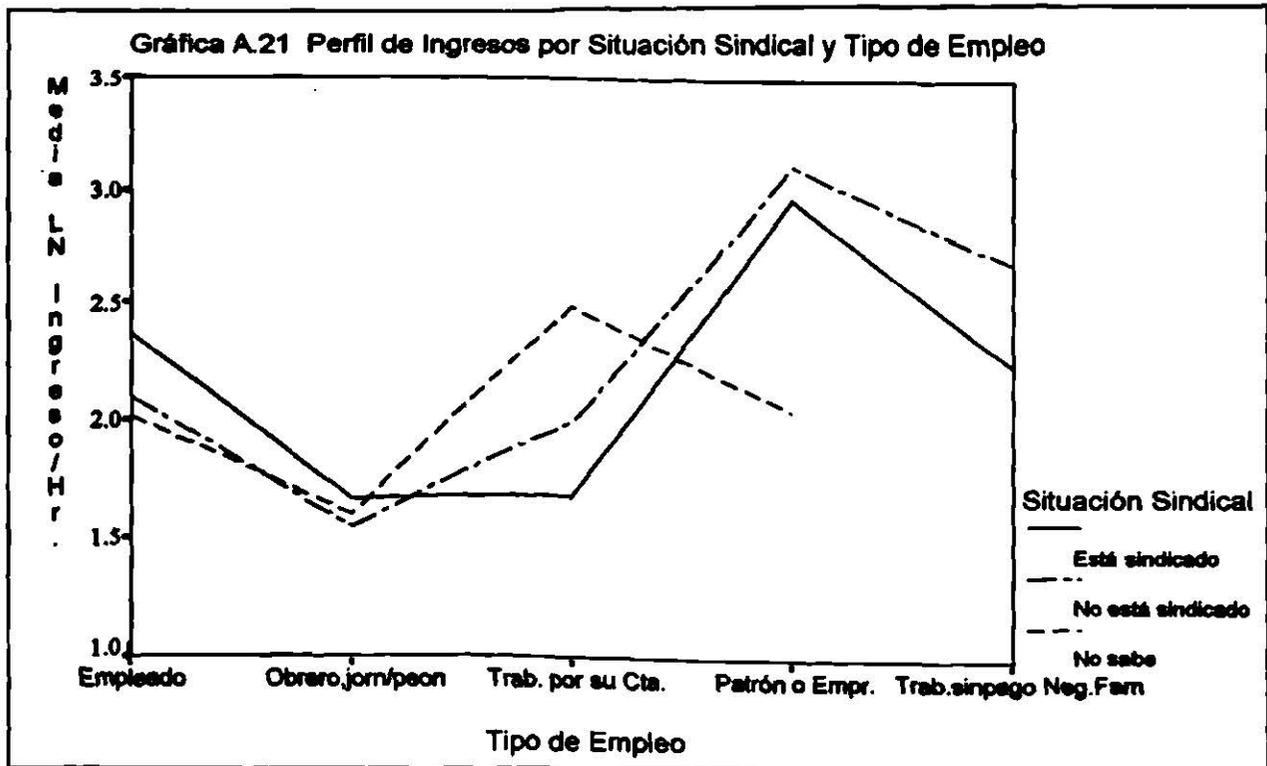
Gráfica A.17 Perfil de Ingresos por Antigüedad y Estado Civil

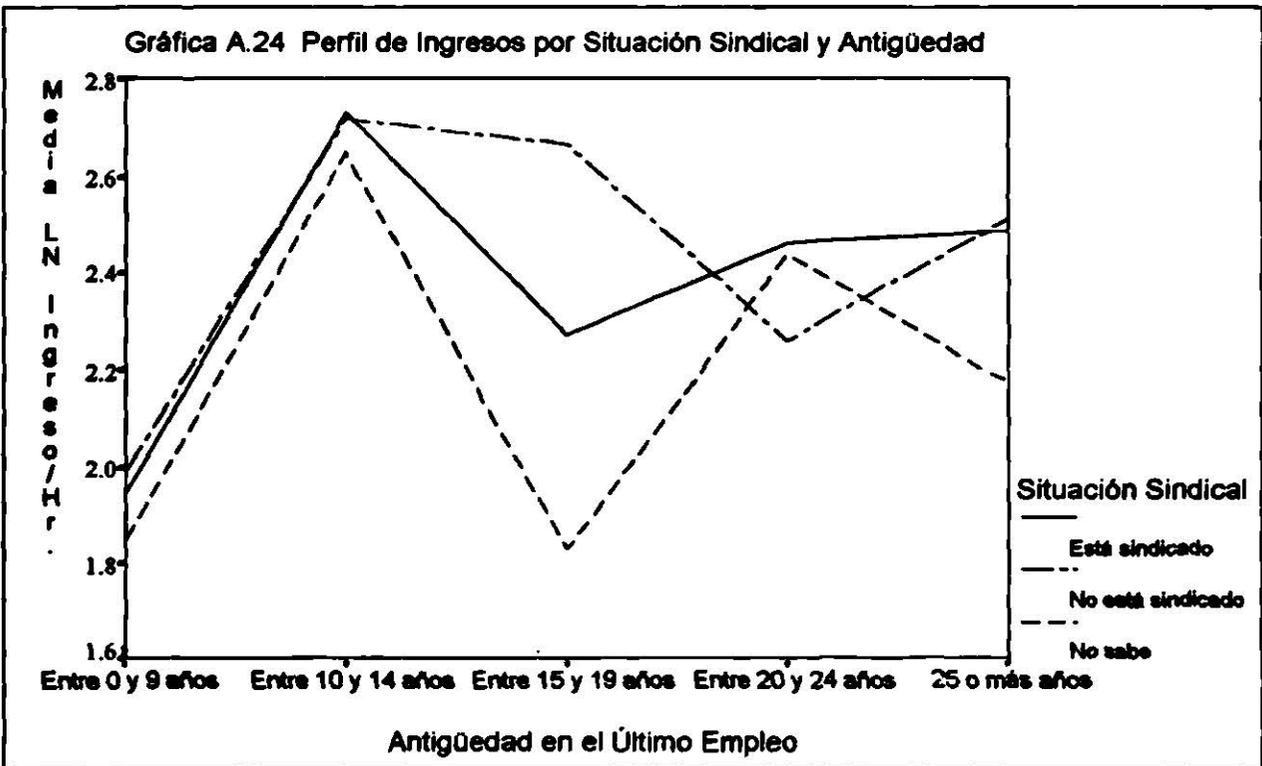
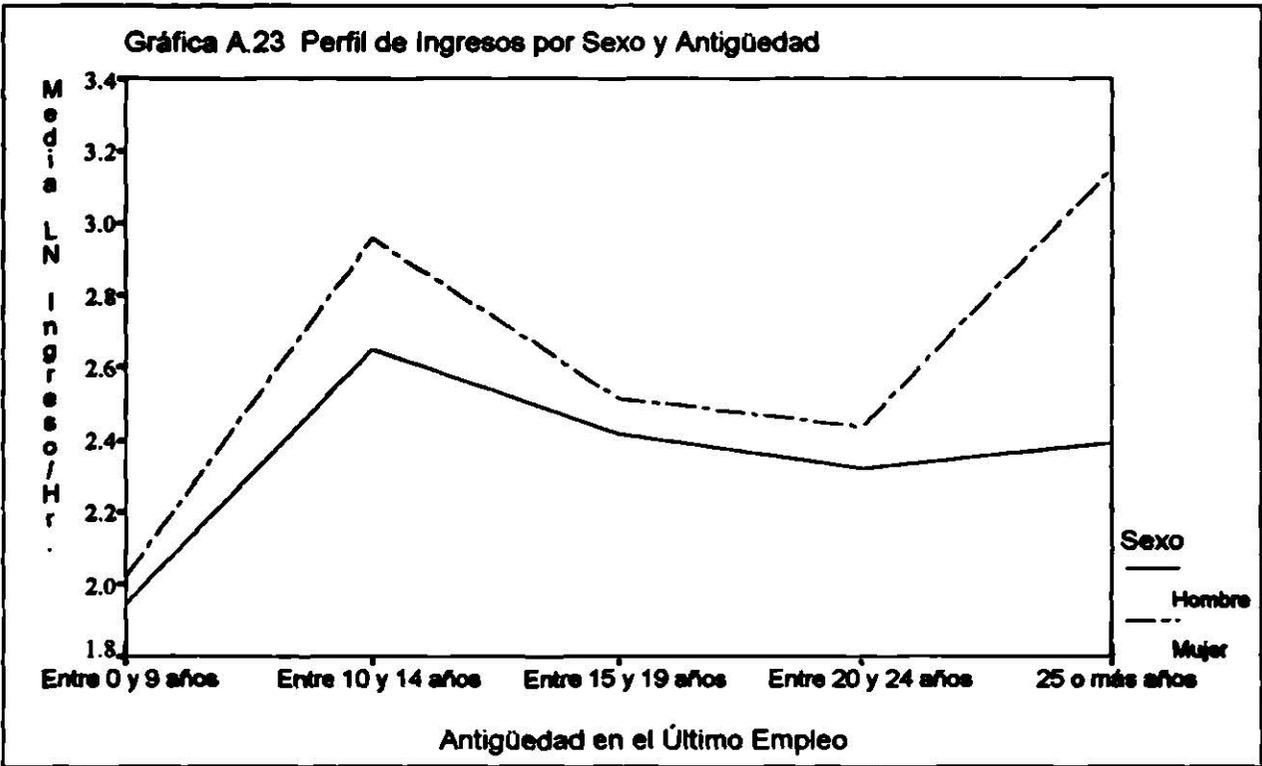


Gráfica A.18 Perfil de Ingresos por Antigüedad y Tipo de Empleo

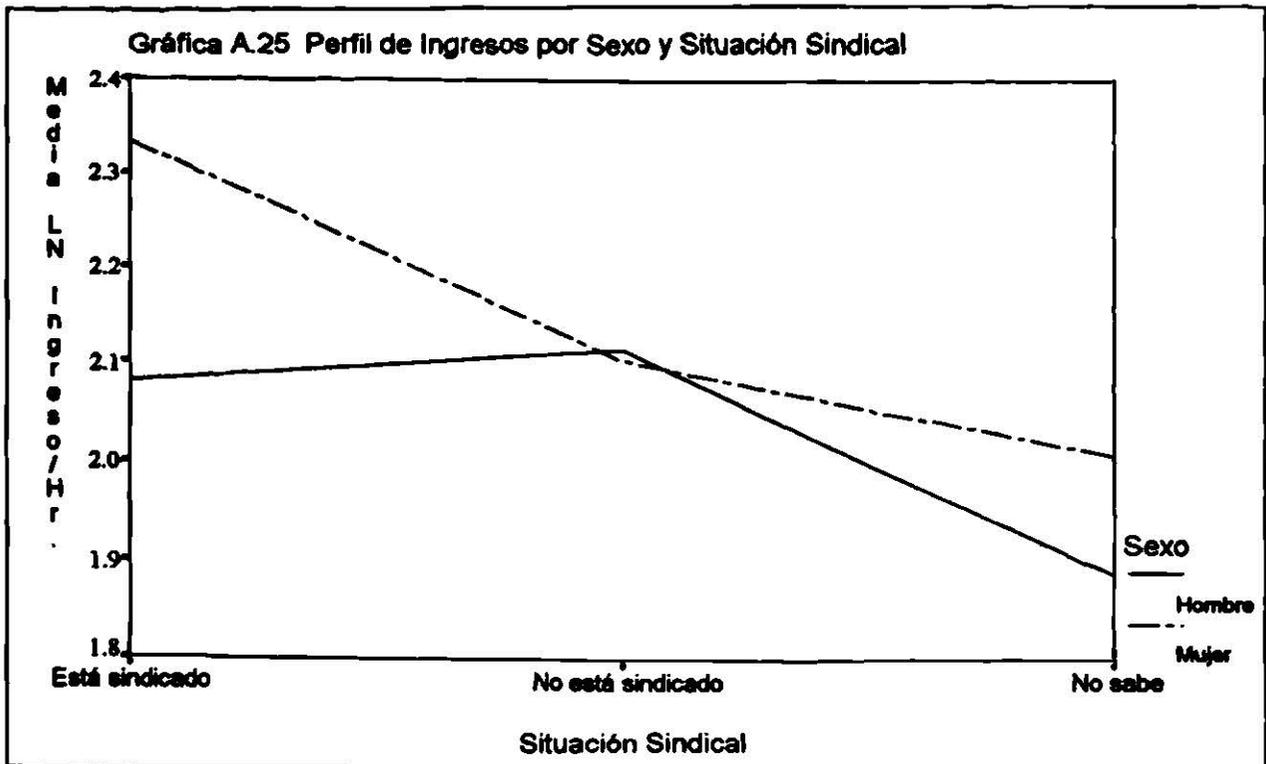




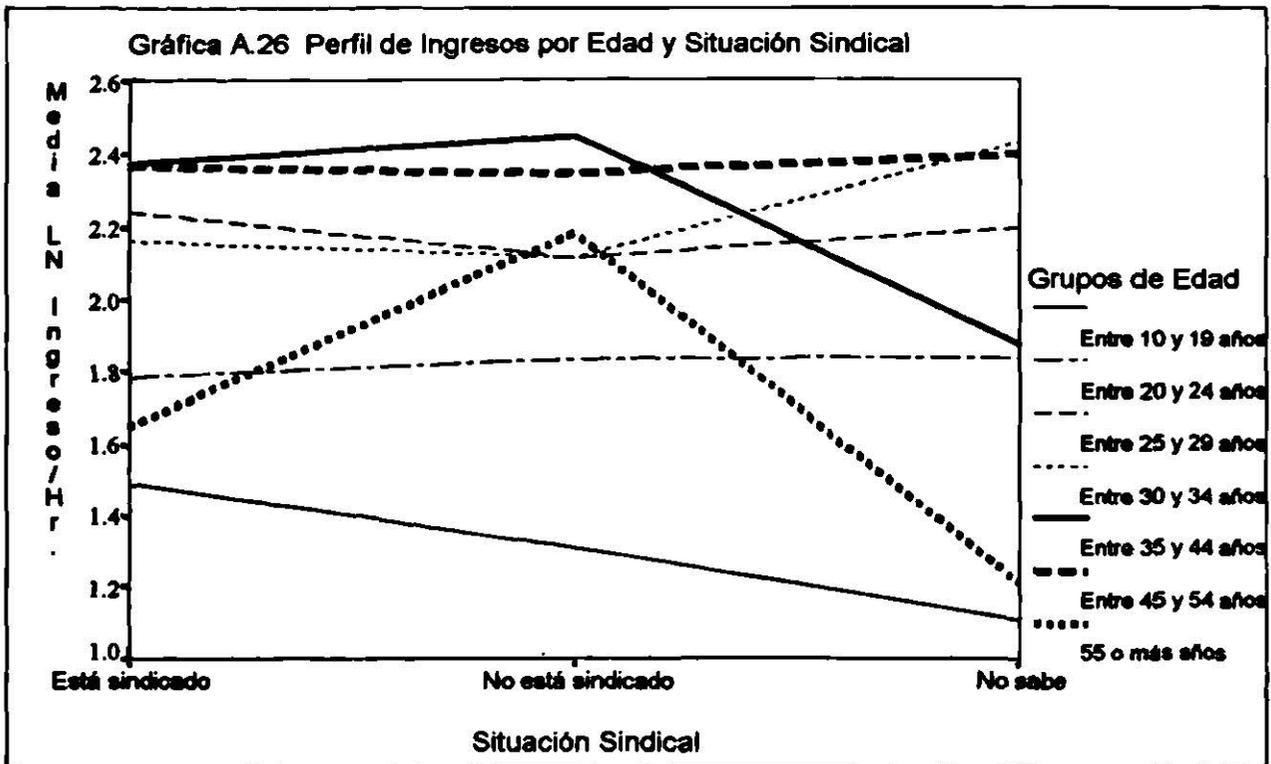




Gráfica A.25 Perfil de Ingresos por Sexo y Situación Sindical



Gráfica A.26 Perfil de Ingresos por Edad y Situación Sindical



Apéndice 2. Resultados de pruebas estadísticas

General Linear Models Procedure Class Level Information

Class	Levels	Values
SEX	2	1 2
AGE	7	1 2 3 4 5 6 7
EDU	8	1 2 3 4 5 6 7 8
MSTAT	3	1 2 3
SECT	12	1 2 3 4 5 6 7 8 9 10 11 12
EXP	5	1 2 3 4 5
SIND	3	1 2 3
JOB	5	1 2 3 4 5

Number of observations in data set = 1065

General Linear Models Procedure

Dependent Variable: LNWAGE

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	492	729.75136126	1.48323447	3.60	0.0001
Error	572	235.89488186	0.41240364		
Corrected Total	1064	965.64624312			
	R-Square	C.V.	Root MSE	LNWAGE Mean	
	0.755713	30.42984	0.64218661	2.11038417	

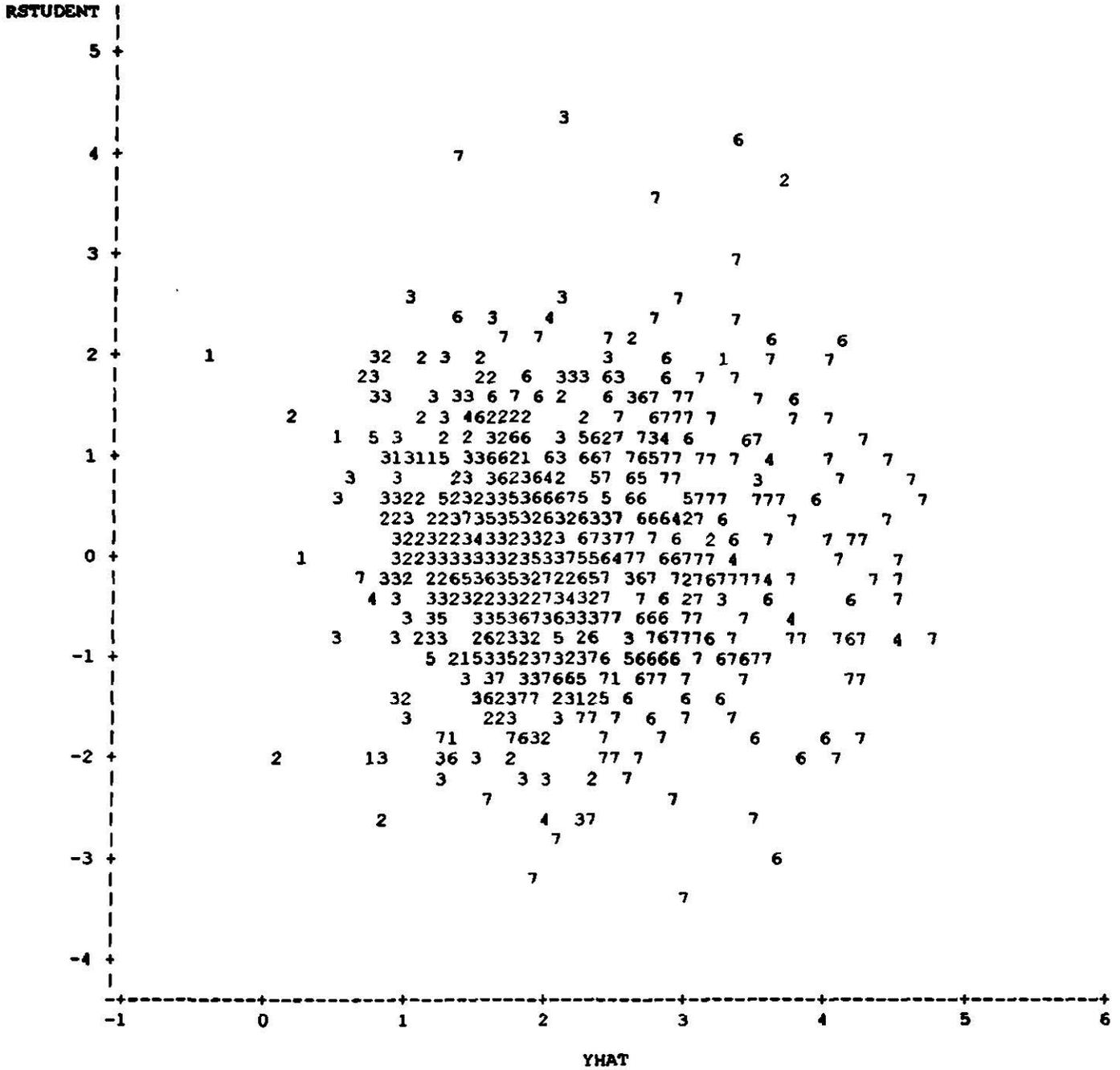
Source	DF	Type III SS	Mean Square	F Value	Pr > F
SEX	1	0.61064029	0.61064029	1.48	0.2242
AGE	6	4.19191923	0.69865321	1.69	0.1201
EDU	7	18.98274443	2.71182063	6.58	0.0001
MSTAT	2	4.28909455	2.14454728	5.20	0.0058
SECT	11	7.31364557	0.66487687	1.61	0.0913
EXP	4	4.97517088	1.24379272	3.02	0.0177
SIND	2	0.63208689	0.31604345	0.77	0.4652
JOB	4	7.59134653	1.89783663	4.60	0.0012
SEX*AGE	6	1.56760974	0.26126829	0.63	0.7035
SEX*EDU	6	2.16672628	0.36112105	0.88	0.5124
SEX*MSTAT	2	0.17972391	0.08986195	0.22	0.8043
SEX*SECT	9	5.75262475	0.63918053	1.55	0.1272
SEX*EXP	4	1.51305191	0.37826298	0.92	0.4535
SEX*SIND	2	0.20091771	0.10045885	0.24	0.7839
SEX*JOB	3	0.41275184	0.13758395	0.33	0.8010
AGE*EDU	33	15.88849107	0.48146943	1.17	0.2420
AGE*MSTAT	11	6.34262735	0.57660249	1.40	0.1694
AGE*SECT	51	24.31527313	0.47677006	1.16	0.2199
AGE*EXP	14	6.63475391	0.47391099	1.15	0.3114

General Linear Models Procedure

Dependent Variable: LNWAGE

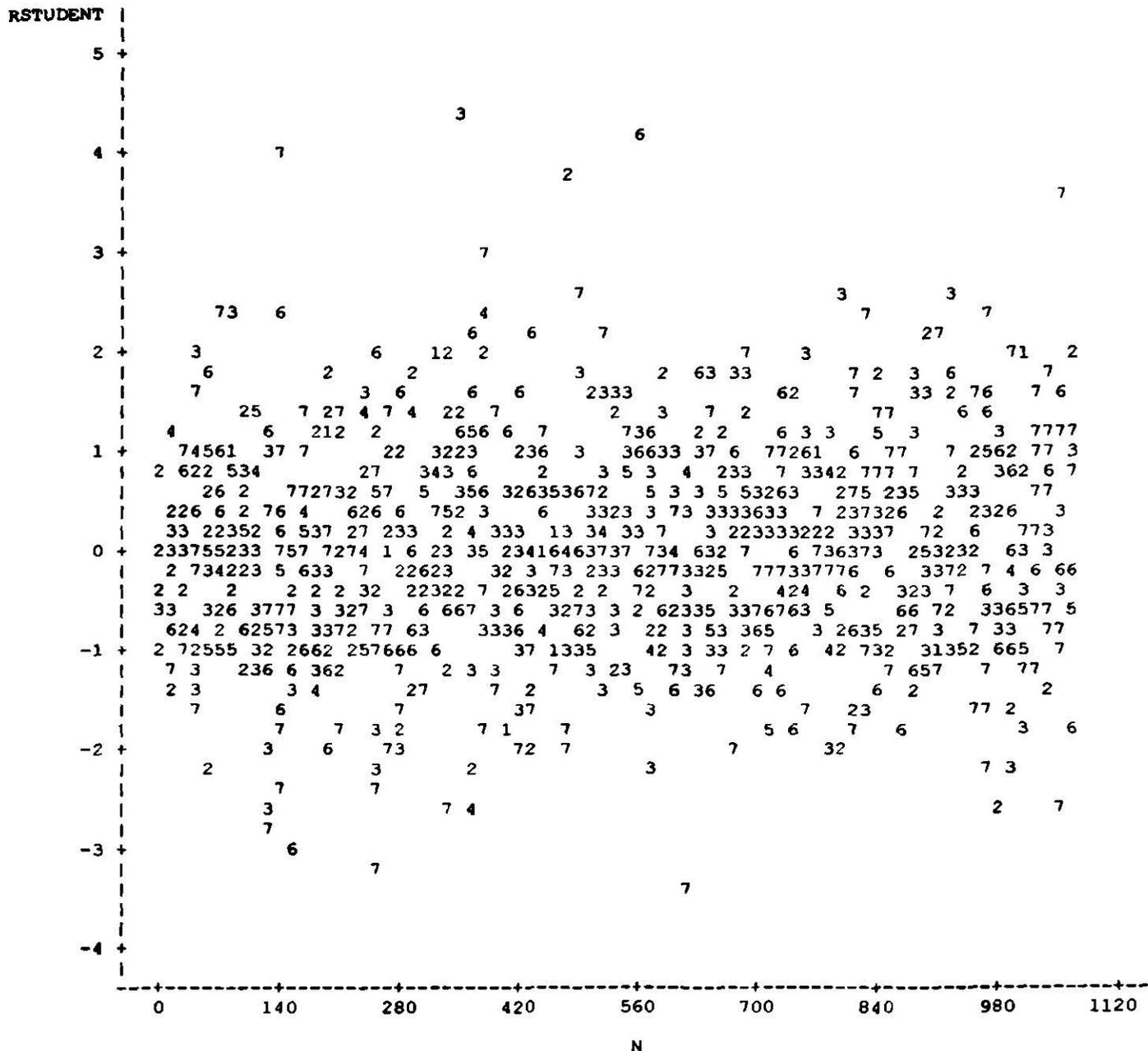
Source	DF	Type III SS	Mean Square	F Value	Pr > F
AGE*SIND	12	4.11256464	0.34271372	0.83	0.6184
AGE*JOB	17	6.42227629	0.37778096	0.92	0.5549
EDU*MSTAT	10	3.45208573	0.34520857	0.84	0.5929
EDU*SECT	49	24.22392214	0.49436576	1.20	0.1733
EDU*EXP	23	12.24211768	0.53226599	1.29	0.1653
EDU*SIND	12	10.83105023	0.90258752	2.19	0.0111
EDU*JOB	18	11.97003082	0.66500171	1.61	0.0521
MSTAT*SECT	15	5.33050810	0.35536721	0.86	0.6080
MSTAT*EXP	7	6.24644546	0.89234935	2.16	0.0358
MSTAT*SIND	4	2.27548242	0.56887061	1.38	0.2397
MSTAT*JOB	6	7.75573341	1.29262224	3.13	0.0050
SECT*EXP	33	19.79347258	0.59980220	1.45	0.0508
SECT*SIND	18	6.19864411	0.34436912	0.84	0.6590
SECT*JOB	23	22.74051836	0.98871819	2.40	0.0003
EXP*SIND	8	3.97358456	0.49669807	1.20	0.2939
EXP*JOB	12	7.18980847	0.59915071	1.45	0.1378
SIND*JOB	6	3.72919505	0.62153251	1.51	0.1734

Plot of RSTUDENT*YHAT. Symbol is value of EDU.



NOTE: 83 obs had missing values. 493 obs hidden.

Plot of RSTUDENT*N. Symbol is value of EDU.



NOTE: 83 obs had missing values. 274 obs hidden.

Levene Test for Equality of Variances for Response Variable LNWAGE
 Use the F-test Statistic for the Overall Model to Test Equality of Variances

General Linear Models Procedure
 Class Level Information

Class	Levels	Values
SEX	2	1 2
AGE	7	1 2 3 4 5 6 7
EDU	8	1 2 3 4 5 6 7 8
MSTAT	3	1 2 3
SECT	12	1 2 3 4 5 6 7 8 9 10 11 12
EXP	5	1 2 3 4 5
SIND	3	1 2 3
JOB	5	1 2 3 4 5

Number of observations in data set = 1065

Levene Test for Equality of Variances for Response Variable LNWAGE
 Use the F-test Statistic for the Overall Model to Test Equality of Variances

General Linear Models Procedure

Dependent Variable: D

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	492	35.22275579	0.07159097	0.99	0.5287
Error	572	41.21493052	0.07205407		
Corrected Total	1064	76.43768630			

R-Square	C.V.	Root MSE	D Mean
0.460804	223.1362	0.26842890	0.12029825

Source	DF	Type I SS	Mean Square	F Value	Pr > F
SEX	1	0.00019282	0.00019282	0.00	0.9588
AGE	6	1.44229600	0.24038267	3.34	0.0031
EDU	7	1.72177779	0.24596826	3.41	0.0014
MSTAT	2	0.19621963	0.09810981	1.36	0.2571
SECT	11	2.07475141	0.18861376	2.62	0.0029
EXP	4	1.93741941	0.48435485	6.72	0.0001
SIND	2	0.37581355	0.18790678	2.61	0.0746
JOB	4	1.02962728	0.25740682	3.57	0.0069
SEX*AGE	6	0.15973718	0.02662286	0.37	0.8984
SEX*EDU	7	0.47981191	0.06854456	0.95	0.4663
SEX*MSTAT	2	0.16038445	0.08019222	1.11	0.3293
SEX*SECT	11	0.65664577	0.05969507	0.83	0.6115
SEX*EXP	4	0.16212007	0.04053002	0.56	0.6900
SEX*SIND	2	0.26194335	0.13097167	1.82	0.1633
SEX*JOB	4	0.41285450	0.10321363	1.43	0.2218
AGE*EDU	38	2.79684014	0.07360106	1.02	0.4368
AGE*MSTAT	11	1.40447288	0.12767935	1.77	0.0556

General Linear Models Procedure

Source	DF	Type I SS	Mean Square	F Value	Pr > F
AGE*SECT	61	4.49810184	0.07373937	1.02	0.4314
AGE*EXP	15	0.47721604	0.03181440	0.44	0.9664
AGE*SIND	12	0.39072594	0.03256049	0.45	0.9415
AGE*JOB	20	1.72677211	0.08633861	1.20	0.2493
EDU*MSTAT	12	0.21446806	0.01787234	0.25	0.9956
EDU*SECT	56	3.50536911	0.06259588	0.87	0.7395
EDU*EXP	26	0.69472791	0.02672030	0.37	0.9984
EDU*SIND	12	0.67770784	0.05647565	0.78	0.6675
EDU*JOB	19	1.53302361	0.08068545	1.12	0.3259
MSTAT*SECT	17	0.87562161	0.05150715	0.71	0.7890
MSTAT*EXP	7	0.19258151	0.02751164	0.38	0.9131
MSTAT*SIND	4	0.04887273	0.01221818	0.17	0.9539
MSTAT*JOB	6	0.22204367	0.03700728	0.51	0.7982
SECT*EXP	36	1.46938426	0.04081623	0.57	0.9813
SECT*SIND	18	0.94282958	0.05237942	0.73	0.7845
SECT*JOB	23	0.99301321	0.04317449	0.60	0.9306
EXP*SIND	8	0.81282442	0.10160305	1.41	0.1890
EXP*JOB	12	0.43198427	0.03599869	0.50	0.9153
SIND*JOB	6	0.24257991	0.04042998	0.56	0.7614

Levene Test for Equality of Variances for Response Variable LNWAGE
 Use the F-test Statistic for the Overall Model to Test Equality of Variances

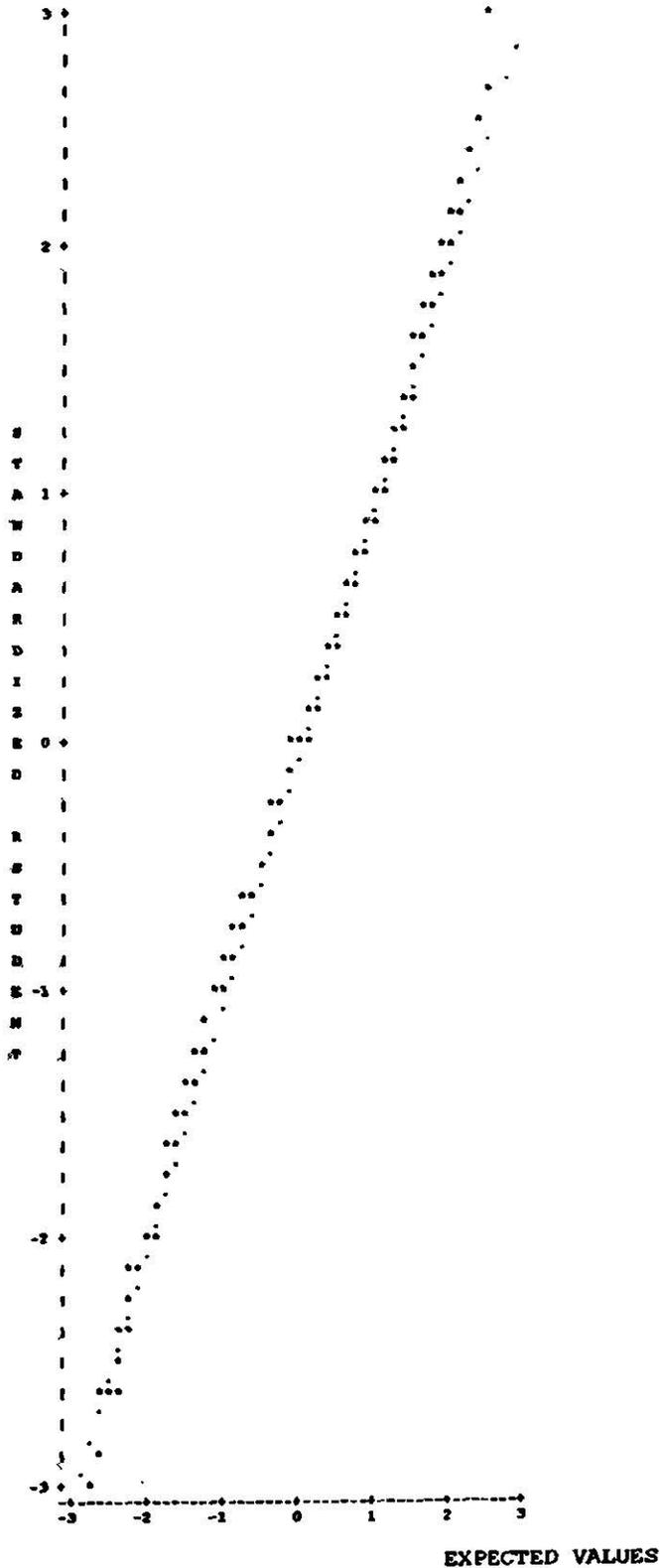
General Linear Models Procedure

Dependent Variable: D

Source	DF	Type III SS	Mean Square	F Value	Pr > F
SEX	1	0.08041245	0.08041245	1.12	0.2912
AGE	6	0.17256904	0.02876151	0.40	0.8797
EDU	7	0.13772667	0.01967524	0.27	0.9643
MSTAT	2	0.01061103	0.00530551	0.07	0.9290
SECT	11	0.41451409	0.03768310	0.52	0.8883
EXP	4	0.02866363	0.00716591	0.10	0.9826
SIND	2	0.04268206	0.02134103	0.30	0.7438
JOB	4	0.07485342	0.01871336	0.26	0.9037
SEX*AGE	6	0.46103021	0.07683837	1.07	0.3815
SEX*EDU	6	0.46478905	0.07746484	1.08	0.3761
SEX*MSTAT	2	0.22356415	0.11178207	1.55	0.2128
SEX*SECT	9	0.63506256	0.07056251	0.98	0.4560
SEX*EXP	4	0.06803191	0.01700798	0.24	0.9180
SEX*SIND	2	0.12181535	0.06090767	0.85	0.4300
SEX*JOB	3	0.20403418	0.06801139	0.94	0.4190
AGE*EDU	33	3.35923907	0.10179512	1.41	0.0655
AGE*MSTAT	11	0.62742301	0.05703846	0.79	0.6487
AGE*SECT	51	3.79846491	0.07447970	1.03	0.4136
AGE*EXP	14	0.43625184	0.03116085	0.43	0.9643
AGE*SIND	12	0.35026081	0.02918840	0.41	0.9618
AGE*JOB	17	0.84485337	0.04969726	0.69	0.8146
EDU*MSTAT	10	0.35485487	0.03548549	0.49	0.8953
EDU*SECT	49	3.36148845	0.06860181	0.95	0.5683
EDU*EXP	23	0.53785510	0.02338500	0.32	0.9990
EDU*SIND	12	0.92949449	0.07745787	1.07	0.3787
EDU*JOB	18	1.15447201	0.06413733	0.89	0.5910
MSTAT*SECT	15	0.54249986	0.03616666	0.50	0.9400
MSTAT*EXP	7	0.27987621	0.03998232	0.55	0.7926
MSTAT*SIND	4	0.02565172	0.00641293	0.09	0.9859
MSTAT*JOB	6	0.22229660	0.03704943	0.51	0.7978
SECT*EXP	33	1.23883982	0.03754060	0.52	0.9882
SECT*SIND	18	1.11895086	0.06216394	0.86	0.6249
SECT*JOB	23	1.01747285	0.04423795	0.61	0.9209
EXP*SIND	8	0.66718415	0.08339802	1.16	0.3230
EXP*JOB	12	0.47400349	0.03950029	0.55	0.8830
SIND*JOB	6	0.24257991	0.04042998	0.56	0.7614

NORMPLT 2.0: INCOME ANOVA NINE FACTORS

Var: RSTUDENT Sample Size: 982 Correlation (obs-vs-exp): 0.9947
Lilliefors's Statistic: 0.0305 Approx P-value: 0.0311
Plot Symbols: '*' for obs-vs-exp and '.' for exp-vs-exp



_NOTE: 166 obs had missing values. 1830 obs hidden. 9 obs out of range.

Apéndice 3. Análisis de Varianza para México, por regiones y nivel educativo

A continuación se presentan los análisis de varianza por regiones y niveles educativos. Nótese que en el caso de las regiones, el factor educativo siempre es el que mayor proporción de la varianza explicada posee y el factor sexo no posee un poder explicativo sustancial. Mientras que en el caso de los niveles educativos, claramente se observa que el factor sexo solamente es estadísticamente significativo en el caso de los niveles básicos de educación, primaria y secundaria, pero deja de serlo para los niveles superiores, preparatoria y licenciatura, es decir, a nivel nacional no se presentan diferencias entre los hombres y las mujeres pertenecientes a la PEA con educación preparatoria y superior, en cuanto a los ingresos por hora se refiere.

Cuadro A.1 ANOVA para la República Mexicana

Fuente de Variación	Proporción de la Varianza Explicada (%)	"F" Marginal^c	Valor de "P"
Último Grado de Estudio Alcanzado	4.610	582.25	0.0001
Estado de la República	3.012	217.10	0.0001
Oficio que Desarrolla	2.994	1114.89	0.0001
Tipo de Empleo	1.985	1108.86	0.0001
Grupos de Edad	0.723	404.04	0.0001
Tipo de Empresa	0.387	172.91	0.0001
Estado civil	0.364	407.49	0.0001
Actividad Económica	0.222	99.34	0.0001
Tamaño de la Empresa	0.165	122.72	0.0001
Sexo	0.077	171.82	0.0001
Covarianza	32.329		
Total Explicado^a	46.868	1562.69	0.0001
Individual^b	53.132		
Total	100		

Muestra = 118,761 Observaciones.

- a. Se refiere a la variación total explicada entre los diversos niveles de los factores.
- b. Variabilidad dentro de los niveles de los factores.
- c. Marginal, por ser calculada para cada factor explicativo y para el total explicado.

Cuadro A.2 ANOVA para el Distrito Federal

Fuente de Variación	Proporción de la Varianza Explicada (%)	"F" Marginal^c	Valor de "P"
Último Grado de			
Estudio Alcanzado	6.783	71.82	0.0001
Oficio que Desarrolla	2.560	27.11	0.0001
Tipo de Empleo	1.969	41.72	0.0001
Grupos de Edad	0.945	15.01	0.0001
Actividad Económica	0.446	5.67	0.0001
Estado civil	0.363	11.52	0.0001
Tipo de Empresa	0.186	2.37	0.0370
Tamaño de la Empresa	0.174	3.67	0.0120
Sexo	0.002	0.15	0.7030
Covarianza	33.634		
Total Explicado^a	47.634	85.42	0.0001
Individual^b	52.937		
Total	100		

Muestra = 3,399 Observaciones.

- Se refiere a la variación total explicada entre los diversos niveles de los factores.
- Variabilidad dentro de los niveles de los factores.
- Marginal, por ser calculada para cada factor explicativo y para el total explicado.

Cuadro A.3 ANOVA para la Zona Metropolitana de Guadalajara, Jalisco

Fuente de Variación	Proporción de la Varianza Explicada (%)	"F" Marginal^c	Valor de "P"
Último Grado de			
Estudio Alcanzado	5.488	54.90	0.0001
Oficio que Desarrolla	3.063	30.64	0.0001
Tipo de Empleo	1.708	34.17	0.0001
Grupos de Edad	1.029	15.45	0.0001
Actividad Económica	0.499	5.99	0.0001
Estado civil	0.449	13.47	0.0001
Sexo	0.412	24.73	0.0001
Tipo de Empresa	0.402	4.830	0.0001
Tamaño de la Empresa	0.125	2.50	0.0580
Covarianza	33.541		
Total Explicado^a	46.718	80.12	0.0001
Individual^b	53.282		
Total	100		

Muestra = 3,234 Observaciones.

- Se refiere a la variación total explicada entre los diversos niveles de los factores.
- Variabilidad dentro de los niveles de los factores.
- Marginal, por ser calculada para cada factor explicativo y para el total explicado.

Cuadro A.4 ANOVA para la República Mexicana por nivel educativo

Fuente de Variación	<u>Licenciatura</u>			<u>Preparatoria</u>		
	Proporción de la Varianza Explicada (%)	"F" Marginal ^c	Valor de "P"	Proporción de la Varianza Explicada (%)	"F" Marginal ^c	Valor de "P"
Oficio que Desarrolla	5.339	190.32	0.0001	6.322	181.48	0.0001
Estado de la República	3.855	26.59	0.0001	4.131	22.95	0.0001
Tipo de Empleo	2.310	123.54	0.0001	4.151	178.74	0.0001
Grupos de Edad	1.331	71.18	0.0001	0.934	40.23	0.0001
Estado civil	0.926	99.08	0.0001	0.644	55.42	0.0001
Tamaño de la Empresa	0.516	36.78	0.0001	0.379	21.79	0.0001
Actividad Económica	0.351	15.04	0.0001	0.604	20.82	0.0001
Tipo de Empresa	0.338	14.44	0.0001	0.320	11.024	0.0001
Sexo	0.018	3.92	0.0480	0.003	0.48	0.4900
Covarianza	10.654			12.985		
Total Explicado ^a	25.639	89.89	0.0001	30.473	86.04	0.0001
Individual ^b	74.361			69.527		
Total	100			100		
Muestra		15,967			12,037	

Fuente de Variación	<u>Secundaria</u>			<u>Primaria</u>		
	Proporción de la Varianza Explicada (%)	"F" Marginal ^c	Valor de "P"	Proporción de la Varianza Explicada (%)	"F" Marginal ^c	Valor de "P"
Estado de la República	5.136	66.89	0.0001	5.249	67.27	0.0001
Oficio que Desarrolla	3.776	254.08	0.0001	2.302	152.42	0.0001
Tipo de Empleo	3.010	303.84	0.0001	3.549	352.46	0.0001
Grupos de Edad	1.520	153.43	0.0001	1.186	117.79	0.0001
Estado civil	0.360	72.77	0.0001	0.289	57.36	0.0001
Tamaño de la Empresa	0.295	39.70	0.0001	0.105	13.928	0.0001
Actividad Económica	0.442	35.72	0.0001	0.405	32.18	0.0001
Tipo de Empresa	0.756	61.04	0.0001	0.588	46.67	0.0001
Sexo	0.168	67.787	0.0001	0.343	36.42	0.0001
Covarianza	14.348			13.405		
Total Explicado ^a	29.812	197.31	0.0001	27.423	178.56	0.0001
Individual ^b	70.188			72.577		
Total	100			100		
Muestra		28,400			28,889	

- Se refiere a la variación total explicada entre los diversos niveles de los factores.
- Variabilidad dentro de los niveles de los factores.
- Marginal, por ser calculada para cada factor explicativo y para el total explicado.

