



**UNIVERSIDAD DE CHILE
FACULTAD DE ECONOMIA Y NEGOCIOS
ESCUELA DE ECONOMIA Y ADMINISTRACION**

DISCRIMINACION POR ETNIA EN EL MERCADO LABORAL CHILENO

**Seminario para optar al Título de
Ingeniería Comercial, Mención Economía**

**Autores:
Francisca Moreno Montealegre
Néstor Romero Navarro**

**Profesor Guía:
Michael Basch Harper**

Santiago de Chile

Julio, 2011

Tabla de contenido

INTRODUCCIÓN.....	3
SECCIÓN I.....	6
REVISIÓN BIBLIOGRÁFICA.....	6
SECCIÓN II.....	11
ANÁLISIS DE DATOS.....	11
Caracterización de la muestra.....	11
SECCIÓN III.....	17
METODOLOGÍA.....	17
3.1 Modelo de estimación por mínimos cuadrados ordinarios.....	17
3.2 Modelo de sesgo de selección, estimación por máxima verosimilitud.....	19
3.3 Descomposición de Blinder – Oaxaca.....	21
3.4 Estadísticos de Autocorrelación Espacial.....	23
3.5 Regresión Espacial.....	24
SECCIÓN IV.....	27
RESULTADOS.....	27
4.1 Resultados para el año 1996.....	27
4.2 Resultados para el año 2000.....	32
4.3 Resultados para el año 2009.....	37
4.4 Resultados descomposición de Blinder – Oaxaca.....	42
4.5 Resultados de los test autocorrelación espacial 2009.....	43
SECCIÓN V.....	44
CONCLUSIONES.....	46
SECCIÓN VII.....	48
REFERENCIAS.....	48
ANEXOS.....	50

INTRODUCCIÓN

El tema de la inclusión indígena en el sector laboral ha sido estudiado de forma exhaustiva a nivel internacional por varios años. A nivel país, tenemos cierta cantidad de estudios acerca de este tema, los cuales han ido aumentando con los años, pero que nos quedamos atrás con países como Estados Unidos y México. Es más, sólo algunos autores del mundo académico se han referido a él durante los últimos años¹, gracias a que se ha ido introduciendo la medición de la variable étnica en las encuestas de interés nacional. En Chile, se ha intentado progresivamente incluir a la etnias dentro de los programas sociales y para ello se conformó la Corporación Nacional de Desarrollo Indígena (CONADI). Bajo esta corporación, se ha fomentado el desarrollo de los pueblos originarios, reconocidos por la ley, en distintos ámbitos tales como: reconocimiento de tierras, nombramiento de monumentos nacionales y establecimiento de áreas de desarrollo indígena. Con respecto a los temas relacionados con el mercado laboral, no se ha avanzado de manera contundente. Es en este pilar en el que los estudios de economía laboral deberían dirigir sus mayores fuerzas.

Nosotros, a través del presente trabajo, intentaremos aportar con datos y mediciones empíricas para que el tema de la discriminación étnica comience a ser parte de la discusión política a nivel nacional y tome la importancia que merece. El tema central de nuestro trabajo es observar la posible discriminación dentro del mercado laboral que pudiese existir hacia las personas con origen étnico. Para ello, contamos con series de datos obtenidos de la encuesta de Caracterización Socioeconómica Nacional (CASEN) para los años 1996, 2000 y 2009². A través del estudio de estos datos hemos podido observar, de forma preliminar, una brecha discriminativa, entre la gente con trabajo y sin trabajo, corrigiendo por el factor. En particular, la gente con origen étnico presenta una menor proporción de gente empleada. Además, dentro de las personas que se encuentran empleadas y que pertenecen a una etnia, un porcentaje mínimo trabaja en empleos de carácter profesionales dentro del mercado formal del trabajo. Un segundo enfoque que pretendemos abordar, es el de la autocorrelación espacial, es decir, estudiar la correlación que puedan tener los datos a través de ciertas localidades, que en nuestro caso serán las provincias. La idea principal de

¹ Dentro de este grupo destacamos al trabajo de Carolina Paz y Alejandra Mizala, del cual haremos referencia más adelante.

² Consideramos estos años ya que desde 1996 se empieza con la pregunta referente a la etnia, posteriormente el 2000 y 2009 se escogen porque representan una buena brecha de tiempo para comparar períodos. Se eliminan los Diaguitas de la muestra.

esto es mostrar una mejora a las estimaciones normales, en donde no se corrige por la correlación que puedan tener ciertos sectores que colindan entre si.

Para intentar explicar esta posible discriminación, primero observaremos características individuales de estas personas, tales como su educación y el lugar dónde viven. Para realizar el estudio, nuestra muestra estará focalizada hacia las etnias con mayor representación dentro del espectro total de pueblos originarios reconocidos por la Ley, que en este caso corresponde a mapuches y aymaras. En general, usaremos variables referentes al capital humano para verificar la existencia de brechas entre las personas con origen étnico y las que no lo poseen. Con los datos obtenidos de la encuesta mencionada anteriormente, seguiremos la estructura de la metodología de Mincer³, en particular, utilizaremos las variables usadas en este tipo de ecuaciones (tales como la educación y la experiencia, por nombrar algunas). Nuestras estimaciones se realizan usando el método de máxima verosimilitud corrigiendo por el sesgo de selección, lo que en la literatura se le conoce como Tobit Generalizado o Tobit II. A diferencia de nuestra metodología, otros trabajos observados, realizan estimaciones de sesgo de selección pero a través de mínimos cuadrados en dos etapas, o también conocido como el método de Heckman en dos Etapas⁴. Nosotros proponemos trabajar con máxima verosimilitud debido a las buenas propiedades que presentan sus estimadores, es decir, estos son consistentes y asintóticamente eficientes. Adicionalmente, realizaremos una descomposición de Blinder – Oaxaca utilizando las variables que resulten ser más significativas en el modelo de sesgo de selección. Con la ayuda de esta descomposición, podremos observar si es que existe realmente una diferencia significativa entre las personas con descendencia étnica y sin esta descendencia. En caso de existir, podremos ver cuánto de ésta se ve explicada por el diferencial en el capital humano y cuánto de ésta se explica por el diferencial en los retornos al capital humano, lo que en la literatura se conoce como discriminación.

Adicionalmente, como se mencionó al principio, pretendemos estudiar la posible existencia de autocorrelación espacial que pueda existir en nuestra muestra. Para ello contamos con las coordenadas tanto de la latitud como de la longitud de las diferentes provincias que componen nuestro país. Con esto queremos agregar un punto de referencia diferente a lo normal, es decir, de encontrar autocorrelación espacial dentro de ciertas provincias, podremos mejorar nuestros

³ CHISWICK, BARRY R., "Jacob Mincer, Experience and the Distribution of Earnings, August 2003

⁴ Heckman, J. 1976. "The common structure of statistical models of truncation, sample selection, and limited dependent variables and a simple estimator for such models". *Annals of Economic and Social Measurement* 5: 475-492

estudios de discriminación étnica. Lo anterior quiere decir que, por ejemplo, el salario de una persona con etnia que reside en la provincia de Cautín puede estar correlacionado con el de otra persona con etnia que reside en la provincia de Malleco. Por lo tanto, presentaremos un precedente con el fin de mejorar las futuras estimaciones que se hagan acerca de la discriminación étnica en nuestro país.

Alguno de los resultados que pudimos observar nos dicen que, en primera instancia, la descendencia étnica afecta el salario en forma negativa (alrededor de un 8%). Sumado a lo anterior, el hecho de pertenecer a ciertas provincias también afecta negativamente el salario, alguna de estas comunas son las que más albergan población indígena, tanto en el norte como en el sur. Por último, según los datos de la CASEN 2009, existe una brecha salarial del 27% entre las personas que tienen origen étnico y las que no lo tienen. De esta brecha, alrededor de la mitad se ve explicado por las diferencias al retorno del capital humano, lo que reflejaría la existencia de una posible discriminación étnica. Con respecto a la autocorrelación espacial, se pudo encontrar que existe una mínima correlación entre las provincias para ciertos factores de interés, como el salario. Si bien es cierto, el valor obtenido de la correlación (positiva) es mínimo, esto puede ser debido a la condición geográfica de nuestro país. Éste está dividido en 53 provincias, dentro de las cuales se puede encontrar poca cercanía entre ellas (debido a la geografía larga y angosta que presenta Chile).

Nuestro trabajo se estructura de la siguiente manera: en la sección I se resumen los principales resultados de trabajos previos realizados tanto a nivel país como a nivel internacional; en la sección II se hace una revisión a la metodología utilizada en nuestro trabajo; en la sección III se presenta un análisis de los datos pertenecientes a la CASEN 1996, 2000 y 2009; en la sección IV se presentan los resultados de la presente entrega; por último en la sección V se resumen las principales conclusiones.

SECCIÓN I

REVISIÓN BIBLIOGRÁFICA.

La discriminación por etnia ha sido objeto de estudio en varios países, dentro de los cuales destaca Estados Unidos. Lo anterior se debe en gran parte a la gran población de raza negra que vive en dicho país, seguido por la inmigración asiática e hispana. Un estudio realizado por Bertrand y Mullainathan (2004) muestra evidencia a favor de la existencia de discriminación por etnia en el mercado laboral de las ciudades de Chicago y Boston. Los autores enviaron CV⁵ ficticios a los avisos de trabajos publicados en los diarios de ambas ciudades. Estos CV fueron contruidos en base a nombres aleatorios de origen africano y blanco. Se reportó que los CV con nombres de origen blanco, recibieron un 50% más de llamadas telefónicas versus los que presenta nombres con origen africano. Además, en términos de CV con un alto potencial, se mantuvo la brecha entre los blancos y afroamericanos. Por otro lado, la característica de residir en un barrio adinerado no ayudaba a los afroamericanos más que a los blancos. Y por último la discriminación hacia los CV con nombres típicos de razas africanas es transversal a todo tipo de industria.

Si bien es cierto, la metodología aplicada en este estudio es de carácter empírico y no tiene un procedimiento econométrico, da cuenta de la existencia en terreno, de ciertos índices de discriminación étnica. Esto nos da una idea de lo que podríamos esperar que suceda para nuestro país. Para ello también se hará un análisis de los datos, sumado a nuestras estimaciones econométricas.

En otro estudio realizado por Reed y Cheng (2003) para el estado de California, Estados Unidos, se encontró evidencia de discriminación salarial negativa hacia hispanos y africanos versus blancos nacidos en dicho país. Sus resultados muestran que para el caso de los hispanos la brecha salarial puede ser explicada por su bajo nivel educacional y status ocupacional. Mientras que para los afroamericanos, su nivel educacional y ocupacional sólo explica una fracción de la brecha salarial. El resto de la brecha no pudo ser explicado por los datos manejados en su trabajo. En resumen, ellos encontraron evidencia de que la brecha salarial, en el estado de California, se ha mantenido en el tiempo, a pesar de que los salarios para los hispanos y afroamericanos han aumentado. Lo anterior se debe a que las políticas públicas, dirigidas a la educación, han permitido un mayor ingreso de este grupo humano a mejores instituciones educacionales, tanto a nivel de secundaria como universitaria. Finalmente, los autores concluyen que mejorando las oportunidades para los

⁵ Curriculum Vitae

trabajadores y las familias de escasos recursos, se podría reducir la brecha salarial en el largo plazo entre los grupos de hispanos y afroamericanos versus los blancos.

Al igual que en el trabajo anterior, este último no presenta un marco de estimación econométrico, simplemente realizan un análisis de datos utilizando la encuesta “Current Population Survey (CPS)”. Esta encuesta se aplica mensualmente a 50 mil hogares de Estados Unidos, y es la primera fuente para conocer las características de la población participante en el mercado laboral. La muestra que utilizaban los autores recoge a empleados entre 25 y 54 años, y para el análisis del año 2000, se basan en datos mensuales de la CPS entre 1999 y 2001. Para el análisis histórico, ellos se basan en los censos aplicados en los años 1980 y 1990.

En relación a Latinoamérica, estudios como el realizado por Atal, Ñopo y Winder (2009) para el Banco Interamericano del Desarrollo, han encontrado evidencia de discriminación salarial por etnia en 18 países pertenecientes a la región, y, en particular, que esta brecha salarial es mayor que la que se puede observar en la diferencia por género. En este trabajo, los investigadores utilizan una técnica no paramétrica de “matching” como una extensión de la descomposición de Blinder-Oaxaca. Sus resultados muestran que las personas pertenecientes a una etnia particular, presentan una brecha mayor en cuanto a logros educacionales, y que esta brecha puede explicar hasta un cuarto de la diferencia salarial total. Se encontró además que, para los trabajos “part-time”, la brecha salarial disminuía entre indígenas y no indígenas, y que la diferencia salarial por etnia está vinculada a la segregación del trabajo. La presencia de personas no pertenecientes a minorías étnicas, que tienen una combinación de características productivas no alcanzable por estas minorías, explica una parte importante de la brecha salarial observada. Aún más, los autores encontraron evidencia, sobre la presencia de barreras en el mercado laboral, que impiden que personas de origen étnico puedan alcanzar puestos importantes en diferentes sectores productivos. Finalmente, los autores proponen mejoras no sólo en el ámbito educacional, sino también hacen hincapié en que se tenga en cuenta lo que el mercado laboral espera de las minorías indígenas. Es decir, a que estas minorías tengan incentivos por parte del mercado laboral para terminar su escolaridad.

En el caso de Chile, se han hecho pocos trabajos relacionados al tema étnico; pero no se ha profundizado tanto como sí se ha hecho para la discriminación por género. Siguiendo esta línea, Jara (2010), en su presentación “Discriminación en mercados laborales por género y etnia en Chile”, analiza ambos focos de discriminación. Para ello, utiliza la encuesta CASEN 2009, en donde

incluye a los trabajadores hombres entre 15 y 65 años y mujeres entre 15 y 60 años, dejando fuera de la muestra, a los trabajadores por cuenta propia y los patrones. En el caso de la discriminación por etnia, el mismo autor encuentra que la diferencia salarial entre mujeres es mayor que en el caso de los hombres. La metodología utilizada por él es la de estimar un modelo de capital humano para caracterizar los ingresos de ambos grupos de trabajadores. Posteriormente, se descompone la brecha salarial según la metodología de Blinder – Oaxaca. Entonces, con esta descomposición, se busca explicar que parte de la brecha corresponde a discriminación, y qué parte se puede explicar por otros factores. Siguiendo este enfoque, se hacen estimaciones independientes para cada grupo por sexo o etnia, corrigiendo por selección. Para ello, se consideran múltiples combinaciones para abarcar todos los cruces, por ejemplo: hombres y mujeres con subgrupos por etnia; indígenas y no indígenas con subgrupos por sexo; comparación entre mujeres indígenas y el resto de la muestra; comparación entre mujeres indígenas y hombres no indígenas.

Para realizar la estimación, el autor utilizó las siguientes variables explicativas: Explicativas cuantitativas: edad, edad al cuadrado, años de escolaridad y escolaridad al cuadrado. Explicativas categóricas: situación de la pareja (la variable toma el valor 0 si no existe pareja o conviviente y toma valor 1 si corresponde a individuos casados o con conviviente); hogar agrícola; reside en la misma comuna hace 5 años; jefe de hogar; rama de actividad; regiones administrativas de Chile.

Los resultados encontrados, dan cuenta de la existencia de discriminación étnica entre indígenas y no indígenas. Se muestra que, a pesar de que la existencia de la brecha, el componente asociado a discriminación es mayor en el caso de los hombres. Para el caso de la mujer indígena, se presenta una mayor diferencia salarial con respecto al hombre no indígena que con el resto de la muestra, a pesar de que el componente no explicado es mayor para la comparación mujer indígena versus resto de la muestra. Respecto al territorio, los resultados de este trabajo muestran que existe una mayor diferencia salarial para el territorio mapuche (provincias de Arauco, Malleco y Cautín) en relación al resto del país.

Por último, quizás el trabajo más completo que se pueda encontrar en la literatura chilena en relación a la discriminación salarial por etnia, es un proyecto de tesis presentado por Carolina Moraga (2008) para optar al grado de Magíster en Economía Aplicada, Universidad de Chile. La

estudiante, quien fue guiada en su trabajo por la profesora Alejandra Mizala, investigó acerca de la posible discriminación que pudiesen estar sufriendo las personas de origen étnico en nuestro país. Para llevar a cabo su proyecto, se utilizó en este trabajo la encuesta CASEN del año 2003, en donde se pudo verificar, a través de los datos, que existe una diferencia salarial, y que los trabajadores indígenas sólo obtenían por su ocupación principal el equivalente al 45% del salario de los trabajadores no indígenas. Es necesario recalcar que la muestra utilizada sólo abarca al grupo de la etnia mapuche, debido a su alta representación dentro de la población. De éstos se eligieron sólo los que habitan en 4 regiones de nuestro país: la octava, la novena, la decima y la Region Metropolitana. A diferencia de las demás etnias reconocidas por la Ley de nuestro país, las personas de origen mapuche superan el 80% de la población indígena total⁶.

Este trabajo, utiliza la metodología de Blinder – Oaxaca, realizando una descomposición por salarios. Para ello, se estiman ecuaciones econométricas de Mincer por Cuantiles, con y sin corrección por sesgo de selección. Los resultados obtenidos muestran que existen diferencias en los retornos al capital humano en contra de los trabajadores mapuches en Chile, y que esta diferencia es mayor entre los hombres. Las mujeres mapuches ganan un 14% menos que las mujeres no indígenas que tienen similares niveles de educación, experiencia y ocupación. Por el lado de los varones, los que pertenecen al grupo de los mapuches ganan un 22% menos que los hombres no indígenas con similares características. La descomposición de Blinder – Oaxaca, revela que las características del capital humano explican un 61% de la diferencia salarial, en tanto que las diferencias en los retornos del capital humano explican casi el 39%. Cuando se corrige por el sesgo de selección (teniendo en cuenta la probabilidad de estar empleado), las diferencias en las características del capital humano explican entre un 52% y un 58% la diferencia de los ingresos; la diferencia en los retornos al capital humano, explica entre un 48% y un 52%, y el sesgo de selección explica entre un -0.1% y -10.5%. Por otro lado, la autora realizó un análisis de las cuantiles, el que permite enfrentar los problemas de heterocedasticidad⁷. Bajo este análisis se encontró que existe dispersión en las diferencias salariales por condición étnica a lo largo de la distribución condicional de ingresos. Además, se encontró que al ser comparados los ingresos de hombres mapuches versus hombres no indígenas, la brecha disminuye hasta el cuantil 0.5 y luego se incrementa nuevamente. En tanto, para el caso de las mujeres la brecha disminuye conforme

⁶ Ver tabla 1 en la sección de Análisis de Datos.

⁷ JONHSTON AND DINARDO, "Econometric Methods" Fourth Edition, Chapter 6: Heteroskedasticity and Autocorrelation

aumentan los ingresos, no existiendo diferencias estadísticamente significativas en los cuantiles más altos. Más allá de lo encontrado por el análisis econométrico, se pudo observar en la muestra que los mapuches tienen menos años de escolaridad y una menor tasa de participación que los no indígenas, ocupándose los primeros mayoritariamente dentro de trabajos no calificados. Además, la autora encontró que en las zonas urbanas y en la Región Metropolitana se perciben mayores ingresos que en las zonas rurales y el resto de las regiones, sin importar el origen étnico o la corrección por sesgo de selección.

Finalmente, los resultados muestran que a través de las distintas metodologías utilizadas, las que no son exactamente idénticas, se puede concluir que los indígenas mapuches reciben ingresos inferiores por su condición étnica. Esto es así dado que el componente de capital humano sólo puede explicar una parte de la brecha, quedando una parte que no es explicada por diferencias de productividad, sino por diferencias en los retornos a ese capital humano.

SECCIÓN II

ANÁLISIS DE DATOS.

Caracterización de la muestra

Para comenzar nuestro análisis de discriminación por etnia, es importante entender cómo se caracteriza la población. Para esto realizaremos inicialmente un análisis descriptivo de la muestra, utilizando datos de la encuesta CASEN 1996, 2000 y 2009.

Debido a cambios en la pregunta de interés a través de los años⁸, se debió evaluar si la muestra variaba significativamente entre un año y otro, ya que en el 2009 no sólo se incorporan como personas de origen étnico a los que se reconocían como tales, sino que adicionalmente se incorporan a las personas que tienen algún grado de ascendencia indígena. Luego de realizar estas comparaciones con los datos, llegamos a la conclusión de que la cantidad de personas no variaba significativamente entre año y año, lo que hace posible realizar comparaciones entre encuestas para diferentes años. Por otra parte, se optó por no considerar a los diaguitas, ya que ellos recién son incluidos en la encuesta del 2009. A continuación, se observa la distribución de las personas entre las distintas etnias.

⁸ Pregunta hecha en la encuesta CASEN 1996: En Chile la ley reconoce la existencia de 8 pueblos originarios o indígenas, pertenece Ud. a alguno de ellos?

Pregunta hecha en la encuesta CASEN 2000: En Chile la ley reconoce la existencia de 8 pueblos originarios o indígenas, pertenece Ud. a alguno de ellos?

Pregunta hecha en la encuesta CASEN 2009: En Chile la ley reconoce la existencia de 9 pueblos indígenas, pertenece Ud. O es descendiente de alguno de ellos?

Tabla 1 Distribución de la población por etnia (en porcentajes)

Etnia	1996	2000	2009
No posee	91,48	91,88	89,44
Aymara	1,45	1,52	1,14
Rapa-nui	0,78	0,01	0,03
Quechua	0,04	0,1	0,13
Mapuche	5,38	6	8,74
Atacameño	0,69	0,39	0,42
Coya	0,04	0,03	0,07
Kawaskar	0,01	0,03	0,02
Yagán	0	0,02	0,02
No sabe	0,13	0	0
TOTAL	100	100	100

Es posible observar que la Ley reconoce la existencia de ocho pueblos originarios, de los cuales los aymara (1,14%) y mapuches (8,74%) representan la mayor parte de la población de indígenas. Para averiguar sobre la participación de cada una de estas etnias dentro de la población, se utiliza la encuesta CASEN, que pregunta sobre la pertenencia a algún pueblo originario. A continuación, es posible observar la cantidad de personas con origen étnico dentro del país, a lo largo de los años.

Tabla 2 Porcentaje de Población con etnia

Indígena	1996	2000	2009
Etnia	8,4	8,11	10,56
sin etnia	91,6	91,89	89,44
TOTAL	100	100	100

Adicionalmente, se puede ver que en el 2009, el 10,56% de la población se reconoce como perteneciente a algún pueblo indígena, siendo que en 1996 eran sólo un 8,4%. Esto puede tener variadas explicaciones, las que se mencionan a continuación: un mayor acceso a la población indígena a través de encuestas; una mayor distribución geográfica de la población dentro del país; un mayor sentido de arraigo hacia las etnias o posibles errores de medición propias de las encuestas.

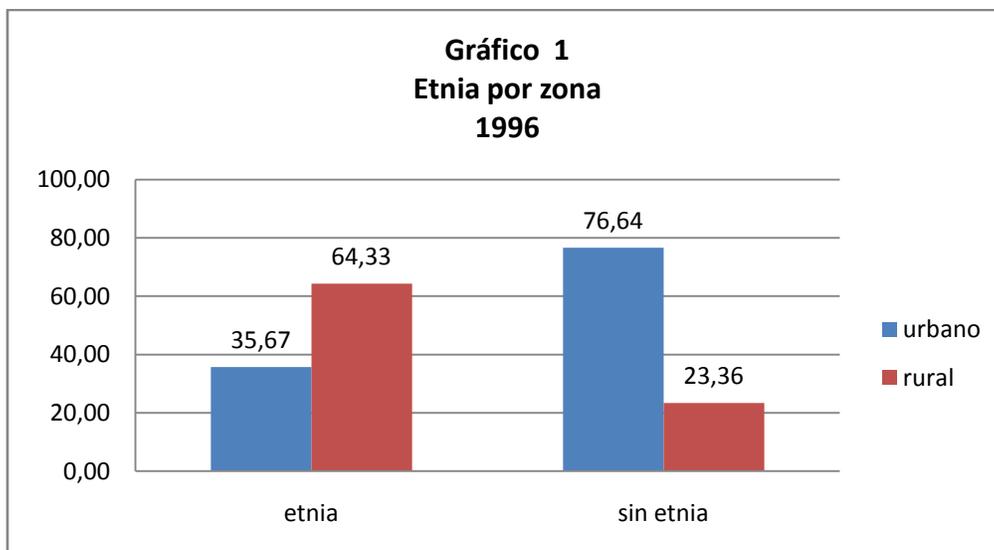
La Tabla 3, nos muestra el porcentaje de personas pertenecientes a alguna etnia, separándolos por género, lo que nos permite ver la participación dentro de la etnia de ambos géneros. Fundamentalmente este ejercicio se realiza para evaluar, que la muestra de pueblos originarios sea símil al resto de la población.

Tabla 3 Porcentaje de personas con etnia por género

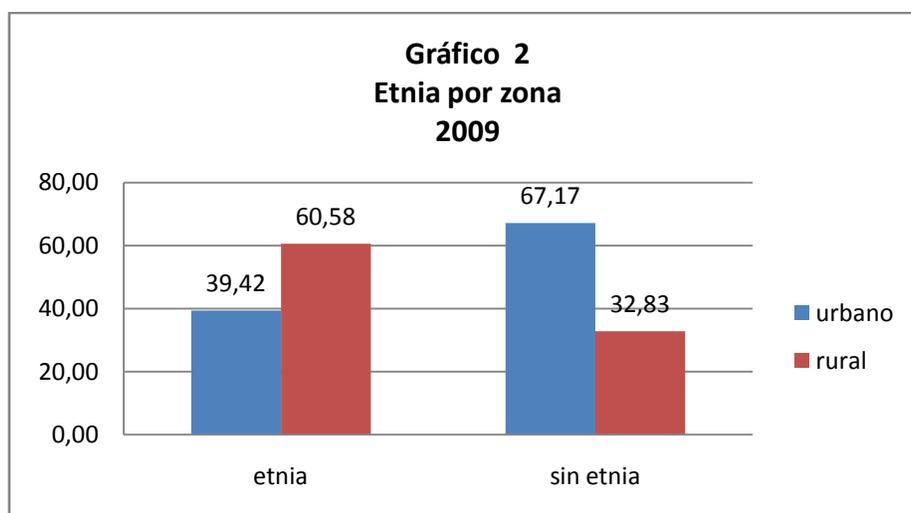
	1996			2000			2009		
	Hombre	mujer	TOTAL	hombre	mujer	TOTAL	hombre	mujer	TOTAL
Etnia	50,69	49,31	100	50,67	49,33	100	49,47	50,53	100
Sin etnia	49,24	50,76	100	49,81	50,19	100	48,99	51,01	100

Observando la muestra es posible notar que para el año 1996 y 2000 se repite el patrón de que es mayor el porcentaje de hombres con etnia que de mujeres, cambiando el sentido en el año 2009, lo que la acerca a los datos de las personas que no se reconocen como de origen indígena.

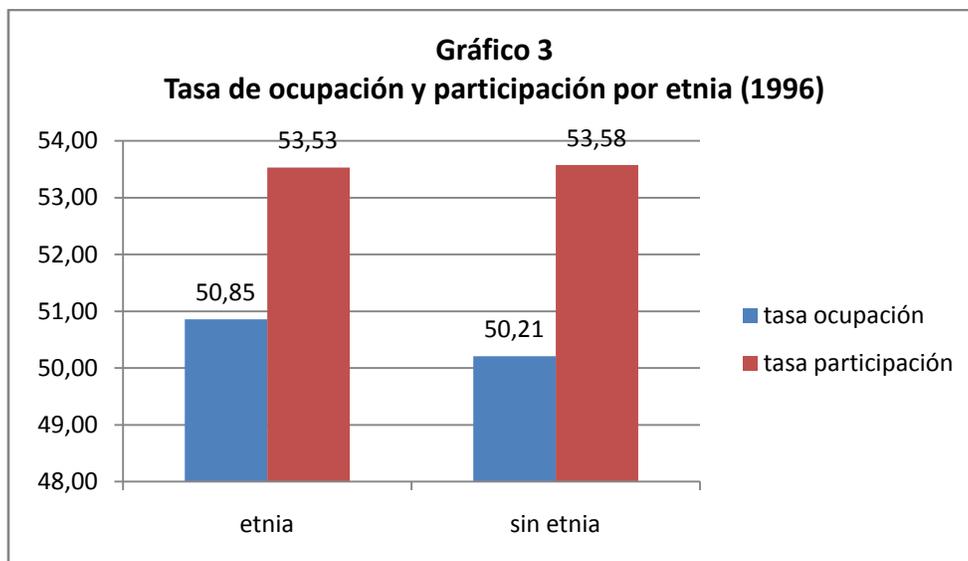
Al evaluar la distribución de la población por zonas en el año 1996, es significativa la diferencia entre los dos grupos, observándose que la población indígena se encuentra mucho más concentrada en zonas rurales (64,33%) que en urbanas, lo que es opuesto en el resto de la población.



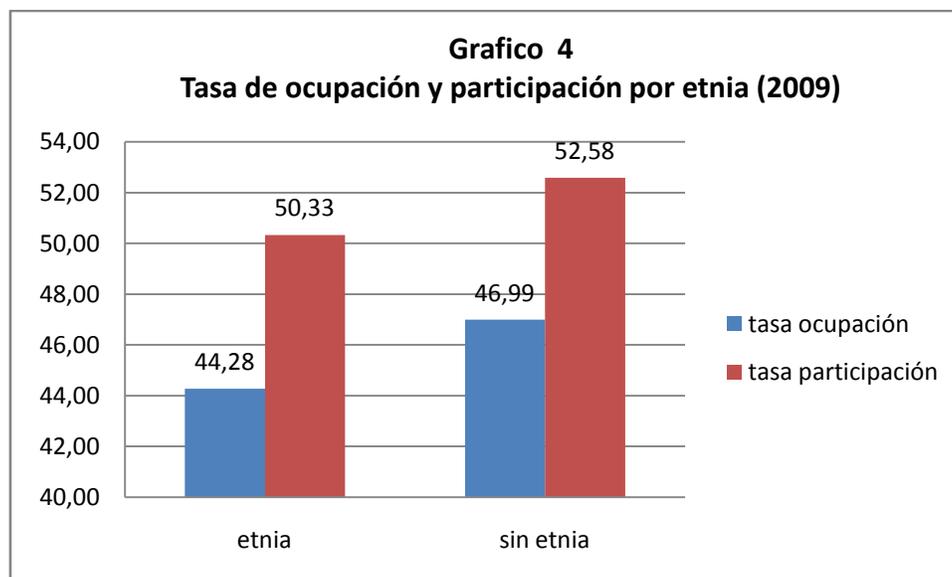
Esta brecha entre zonas rurales y urbanas parece disminuir en el año 2009, aumentando las personas con etnia que pasan al sector urbano y observándose un efecto contrario en el resto de la población. Aún así, es significativa la diferencia entre los dos grupos, donde tan sólo un 39,42% de las indígenas se encuentra en sectores urbanos versus un 67,17% que se observa para el resto.



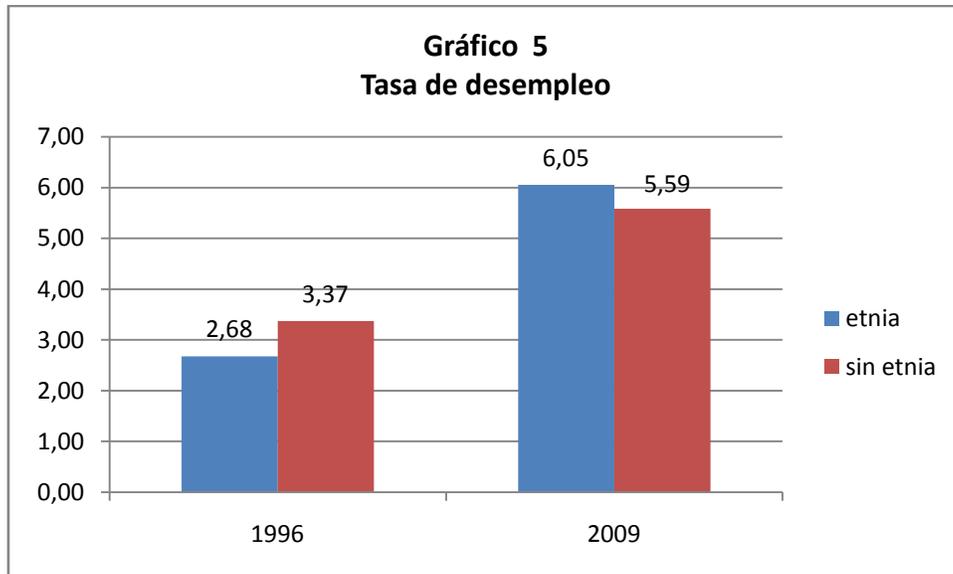
Otra área importante a considerar, es el mercado laboral, para lo cual analizaremos las diferencias entre tasas referentes a la fuerza de trabajo, con el fin de verificar si existe alguna diferencia entre las dos submuestras. Esto nos permitirá hacernos una idea previa de si existe discriminación laboral. En caso de existir, habrá que ver si la discriminación se concentra en la contratación o en la calidad del empleo.



En cuanto a la tasa de participación y ocupación, es posible ver en el gráfico Nro. 3, que en el año 1996 las tasas eran relativamente similares entre los dos grupos, siendo levemente inferior la tasa de ocupación en el caso de las personas sin etnia. Ahora, evaluando para el 2009, el escenario parece cambiar, siendo significativamente menores ambas tasas para los indígenas (ver gráfico Nro. 4).



Si bien estas tasas nos dan una leve idea del mercado laboral en términos de etnia, no nos queda claro que el hecho de tener altas tasas de ocupación asegure que la calidad de empleo sea buena, ya que tal vez la brecha se encuentra ligada a variables vinculadas a la calidad como son contratos estables y protección social, o variables cuantitativas como el ingreso.



Evaluando ahora la tasa de desempleo, parece ser que la situación ha empeorado entre 1996 y el 2009, subiendo significativamente las tasas para ambos grupos, pero en particular para la población indígena, observándose una tasa de 6,05% en el año 2009 versus un 2,68% en 1996.

SECCIÓN III

METODOLOGÍA.

Como se mencionó anteriormente, en esta sección analizaremos la metodología que utilizaremos en nuestra investigación. En primer lugar, presentaremos el modelo lineal que usaremos, las variables dependientes e independientes. En segundo lugar, introduciremos el modelo de sesgo de selección estimado por máxima verosimilitud, modelo que fue construido por Heckman⁹. Este modelo es uno de datos censurados, debido a que no todos los individuos participan del mercado laboral. Por esta razón no podemos observar un valor del salario para todos los individuos, sin embargo sí podemos observar el resto de las variables en ellos. Es por lo anterior, que se estima el modelo usando máxima verosimilitud en conjunto con una ecuación de participación. Luego, realizaremos una descomposición de Blinder – Oaxaca, utilizando las variables más relevantes según lo que hayamos encontrado en los modelos anteriores. Por último, estudiaremos la posibilidad de autocorrelación espacial entre provincias, pudiendo incorporar; en caso de existir; la matriz de pesos espaciales dentro de la estimación.

3.1 Modelo de estimación por mínimos cuadrados ordinarios.

La idea de realizar una estimación utilizando ésta técnica es la de presentar una base de comparación para el modelo de sesgo de selección por máxima verosimilitud. Ambos presentan las mismas variables, sin embargo, el último se estima en conjunto con una ecuación de Participación. El modelo de mínimos cuadrados ordinarios es el siguiente:

La variable dependiente de nuestro modelo es el logaritmo natural del salario. Dentro de las variables, se puede observar que trabajamos con una medida de la experiencia del individuo ($\ln(w_i)$), calculada acorde a la literatura como: la edad del individuo menos sus años de escolaridad menos seis años. La variable w_i^2 corresponde al valor al cuadrado de la experiencia del individuo. Por supuesto, se debería esperar que a mayor experiencia del individuo, mayor debiera ser su salario.

⁹ Heckman, J. 1976. “ The common structure of statistical models of truncation, sample selection, and limited dependent variables and a simple estimator for such models”. Annals of Economic and Social Measurement 5: 475-492.

La principal variable de interés de nuestro análisis, corresponde a que es una variable dummy que toma valor uno si el individuo tiene al menos descendencia indígena (en este caso, puede ser indígena o puede tener descendencia indígena) y cero en el caso contrario. A través de esta dummy esperamos testear nuestra hipótesis, de que existen brechas salariales asociadas al factor étnico, es decir que afecte negativamente el valor del salario.

Por otro lado, agregamos la dummy que hace referencia al género de la persona (, que toma valor uno si la persona pertenece al género masculino, y toma valor cero en caso contrario. De acuerdo a la literatura estudiada, esta variable debería afectar positivamente el nivel del salario. Adicionalmente, se agrega otra dummy que refleja si la persona vive en un sector rural. En este caso la variable toma valor uno si la persona vive en un sector urbano y cero en caso contrario.

Luego, agregamos un set de variables binarias que corresponden al nivel educacional del individuo. En este caso, las variables dan cuenta del último nivel educacional terminado: primario, secundario y terciario. Se debería esperar que a mayor sea el nivel educacional completado, mayor debiese ser el nivel del salario. Por otro lado, se agrega otro set de variables dummy, que informan si el individuo reside en cierta provincia o no. En este caso, contamos con las provincias de Arica, de Parinacota, de la Araucanía, de Cautín, de Malleco, de Chiloé, de Osorno, de Valdivia y por último de Santiago, la razón del uso en particular de estas provincias, se debe a que es allí donde se concentra la población indígena del país. Tanto en las tres provincias del norte como en las cinco del sur, es donde más residen aymara y mapuches, respectivamente. Se espera que la residencia en estas comunas afecten negativamente el nivel del salario, aun que esto quizás tenga más relación con el tema de centralización del país que del problema de discriminación étnico.

Por último, agregamos un set de variables cualitativas que caracterizan al individuo según la rama económica en la que se desempeña. En este caso, las ramas que utilizamos son: agricultura, minería, industrial, gas-agua-electricidad, construcción, comercio, transporte, sector financiero y servicios. Acerca del comportamiento de esta variable, los resultados esperados no son muy claros, es decir podríamos esperar mayores sueldos para el sector financiero y por supuesto en la minería, y quizás sueldos más bajos para el sector agricultura y transporte, pero no optamos por tomar una posición tajante como lo hicimos con las variables anteriores.

3.2 Modelo de sesgo de selección, estimación por máxima verosimilitud.

Además de la metodología MCO, realizamos una estimación controlando por sesgo de selección. En este caso, utilizamos la propuesta desarrollada por Heckman, utilizando el método de máxima verosimilitud. La metodología de Heckman nos permite tener en cuenta el sesgo de selección que se produce cuando uno realiza estimaciones, principalmente las que se realizan sobre el ingreso de las personas¹⁰. En primera instancia lo que se hizo es definir una ecuación de participación, en la cual la variable dependiente es una variable latente que está asociada con la decisión del individuo de participar en el mercado laboral. Por otro lado, las variables que utilizamos como regresores corresponden a variables que, a nuestro juicio, definen que una persona se encuentre o no participando en el mercado laboral. Dentro de estas variables, se encuentran una serie de variables cualitativas que definen aspectos tales como: si la persona tiene descendencia étnica, si la persona reside en un sector urbano, si la persona pertenece al género masculino, si la persona reside en la provincia de Arauco o de Cautín, y si la persona es jefe de su hogar. Adicionalmente, se cuenta con otras variables de carácter cuantitativo tales como: el número de hijos, el número de personas dentro del hogar y las variables que hacen alusión a la experiencia.

Posteriormente, se define la ecuación que estimará el salario de cada individuo. Para ello tenemos nuestra variable dependiente, la cual es el logaritmo del salario del individuo. Y, naturalmente, definimos nuestro set de variables explicativas, las cuales son las mismas ocupadas en el modelo de mínimos cuadrados ordinarios, presentado en el apartado anterior.

Entonces, para comenzar con nuestra estimación debemos especificar el modelo en conjunto, es decir lo que se presenta a continuación:

En este caso, μ representa la variable latente que hace alusión a la decisión del individuo de participar en el mercado laboral. Además, X es un vector de variables determinantes de ésta decisión, las cuales fueron resumidas en párrafos anteriores. La variable ϵ , corresponde al

¹⁰Sesgo de selección: se genera porque la medición de la variable de interés fue truncada de forma no aleatoria. También se le conoce como “truncamiento incidental”.

logaritmo natural del salario y el vector β corresponde a las variables determinantes del salario (las cuales fueron presentados en el modelo MCO en páginas anteriores).

Por lo tanto, se tienen dos grupos de variables. Aquéllas donde y_i (que en nuestro caso, corresponde al valor del salario) es observado, por lo que, sabemos que ϵ_i (que en nuestro caso corresponde a la variable latente) se cumple. Para estas observaciones, la función de verosimilitud es la probabilidad del evento y_i y que también ocurra



Para aquellos valores de y_i no observados, sabemos que $y_i^* < \tau$ y por ende no tenemos información independiente para ϵ_i . Es decir, no observamos el valor del salario de la persona debido a que ésta no se encuentra participando del mercado laboral, pero sí podemos observar sus datos referentes a las demás variables utilizadas:

Juntando ambos términos, tenemos que la estimación es:

Donde existen n_1 observaciones para las cuales no observamos x_1 y n_2 observaciones, donde sí lo hacemos, con n_2 . Se puede observar que, utilizando esta metodología se estiman en conjunto la ecuación de participación y el logaritmo natural del salario del individuo.

3.3 Descomposición de Blinder – Oaxaca.

Esta metodología nos permite observar las diferencias que pueden existir entre dos grupos de una misma muestra. En nuestro caso, en un grupo estarían todos los que tienen descendencia indígena, y en el otro grupo, encontraríamos a las demás personas de la muestra. Entonces, lo que queremos saber, es cuánto de la diferencia entre estos grupos puede ser explicada por ciertas variables de capital humano como la educación, la experiencia laboral, y otras variables relevantes escogidas por nosotros, y cuánto de esta diferencia no puede ser explicado por estas mismas variables, por lo que, a este tramo, se lo puede asignar a la existencia de discriminación, en nuestro caso, discriminación por etnia. Lo anterior, en estricto rigor, al existir una parte no explicada, no necesariamente tiene que asociarse a discriminación. Podríamos decir que no pudo ser explicada por las variables utilizadas, tanto de capital humano como las otras escogidas. Sin embargo, según la literatura, se señala que gran parte de la diferencia de los retornos de capital humano corresponde a la existencia de discriminación.

La metodología de descomposición de Blinder - Oaxaca es la siguiente:

Se tienen dos grupos, el grupo A y el grupo B

Donde y_i , representa el valor esperado de la variable a explicar. Basado en el siguiente modelo lineal:

Donde β representa el vector de variables explicativas, β representa la pendiente de los parámetros y el intercepto y ϵ es una medida del error. La diferencia de las medias de la variable dependiente, puede ser representada como la diferencia de las medias de las variables independientes de cada grupo, esto es:

Ahora la descomposición de Blinder – Oaxaca resulta del hecho de que existe un vector de coeficientes no discriminatorios, que debería ser usado para determinar la contribución en las diferencias de los predictores. Entonces, sea β el vector de coeficientes no discriminatorios. La diferencia resultante puede ser escrita como:

El cual puede ser representado de la siguiente manera:

Donde el primer componente se refiere a:

Ésta es la parte de la diferencia resultante, que es “explicada”, por las diferencias en el capital humano. El segundo componente, se desglosa de la siguiente manera:

el cual corresponde a la parte “no explicada”. Usualmente, es atribuible a la presencia de discriminación, pero también es importante reconocer que esta parte captura todos los efectos potenciales que no son explicados por las variables independientes. En definitiva, se dice que esta parte corresponde a las diferencias en los retornos al capital humano.

En el caso de nuestro modelo, las variables explicativas utilizadas fueron las siguientes: las variables referentes a la experiencia de cada individuo; una dummy referente a si el individuo pertenece al género masculino; tres variables dummy correspondientes al nivel educación de cada individuo, en donde la primera es si el individuo tiene educación secundaria completa, la segunda es si el individuo posee educación secundaria incompleta y la tercera es si el individuo posee educación terciaria completa; tres variables dummy referentes a provincias donde con una alta

población indígena, en este caso son las comunas de Santiago, Malleco y Cautín; por último una variable dummy referente a una rama de actividad económica en particular, el sector de la minería. Todas estas variables fueron escogidas según su nivel de significancia observado en la estimación del modelo de sesgo de selección, explicado en la sección anterior.

3.4 Estadístico de Autocorrelación Espacial

I de Moran

Este estadístico nos ayudará a observar la posible existencia de autocorrelación espacial que pudiese haber entre las provincias de nuestro país. Como se mencionó anteriormente, utilizaremos 53 provincias del país, dejando afuera a Isla de Pascua, pues ésta se encuentra demasiado lejos de Chile continental. Este estadístico da como resultado valores cercanos a cero si no existe autocorrelación espacial, valores cercanos a uno si hay autocorrelación espacial positiva, y valores cercanos a menos uno si autocorrelación espacial negativa¹¹.

Ahora bien, antes de presentar la metodología de este estadístico, es necesario entregar una pequeña explicación de lo que significa autocorrelación espacial. En palabras simples, la autocorrelación espacial significa que, variables correspondientes a una zona en particular (en nuestro caso, a una provincia), pueden estar correlacionadas con las mismas variables de otra zona adyacente, es decir, pueden estar correlacionadas con una zona colindante a la zona inicial. Por ejemplo, podemos decir que el salario de las personas de Malleco, está correlacionado con el salario de las personas de Cautín. Entonces, de existir lo anterior, podríamos decir que estamos en presencia de autocorrelación en series de corte transversal (recordemos que generalmente la autocorrelación es propio de series de tiempo).

A continuación se presenta la metodología a utilizar:

La I de Moran se define como:

¹¹ LeSAGE, JAME P., "Spatial Econometrics", Chapter one, December 1998.

Donde n es el número de unidades espaciales indexadas a i y j ; y_i es la variable de interés; \bar{y} es la media de y ; y w_{ij} es un elemento de la matriz de pesos espaciales. El valor esperado de la I de Moran bajo la hipótesis nula de que no existe autocorrelación espacial es:

Se hace hincapié que la muestra que se utilizará, para obtener este estadístico, se ha visto reducida debido a las limitaciones del software ocupado. Para este caso, sólo se ocuparan 717 observaciones, de las cuales todas corresponden a valores positivos del salario. Lo anterior, corresponde a una muestra censurada, ya que la metodología de la I de Moran no nos permite trabajar con valores no reportados del salario. Esto mismo aplica para el siguiente estadístico.

C de Geary

Éste es otro estadístico que nos permite testear la posible existencia de autocorrelación espacial, que pudiese haber entre las provincias de nuestro país. Este estadístico toma valores entre 0 y 2, donde valores mayores que uno indican la presencia de autocorrelación espacial positiva y menores que uno presencia de autocorrelación negativa¹².

El test de Geary está inversamente relacionado con el de Moran, pero no es idéntico, ya que esta última mide la autocorrelación espacial global, mientras que el test de Geary, mide la autocorrelación parcial. A continuación se presenta la metodología a utilizar:

La C de Geary se define como:

Donde n es el número de unidades espaciales indexadas a i y j ; y_i es la variable de interés; \bar{y} es la media de y ; y w_{ij} es un elemento de la matriz de pesos espaciales¹³.

3.5 Regresión Espacial

Este método corrige por la posible presencia de autocorrelación espacial. En este contexto, a la estimación tradicional¹⁴, se le incorpora un término adicional, que se compone de un vector de

¹² P. DE JONG, C. SPRENGER, AND F VAN VEEN, "On extreme values of Moran's I and Geary's C".

¹³ LeSAGE, JAME P., "Spatial Econometrics", Chapter one, December 1998.

¹⁴ Por ejemplo, el modelo, $y = \beta_0 + \beta_1 x + \epsilon$, que podríamos estimar a través de MCO.

coeficientes rho, multiplicando a una matriz de pesos espaciales y el vector de la variable dependiente . La idea de esta metodología, es no sólo observar cómo afecta un grupo de variables explicativas a una cierta variable dependiente, sino también, incorporar la dependencia de esta última variable con respecto a su similar en el espacio. Lo anterior, se puede explicar, con un ejemplo: podríamos querer observar cómo se ve afectado el salario de la provincia “A” con respecto a ciertas variables explicativas, pero también podríamos observar si el salario de la provincia “B” también afecta el salario de la primera provincia. Para lograr lo anterior, necesitamos una matriz de pesos espaciales. Esta matriz se genera a partir de coordenadas geográficas de los lugares a analizar, tanto su latitud como su longitud, en nuestro caso, las coordenadas de las provincias.

Nuestro trabajo consistió en “geo-referenciar” la CASEN 2009, de tal manera que pudiésemos tener asociadas las variables explicativas a las coordenadas de cada provincia. Para realizar esto utilizamos las coordenadas de cada provincia, sin embargo, debido a que estos datos aparecen asociados a todos los puntos que contiene cada provincia, fue necesario asignar un promedio de las coordenadas a cada una de ellas¹⁵. A continuación se presenta la estructura de la metodología utilizada:

Primero tenemos que construir nuestra matriz W de pesos espaciales, la cual se genera a partir de las coordenadas obtenidas. Una vez ingresadas las coordenadas, nuestra matriz será una matriz cuadrada binaria de NxN, parecida a:

donde los ceros indican que las provincias no son colindantes, y los “unos” corresponden a provincias adyacentes. Ahora bien, lo siguiente en nuestra metodología, es presentar la ecuación a estimar por el método de máxima verosimilitud, la cual se define de la siguiente manera:

¹⁵ Esto podría generar problemas, debido a que la forma de las provincias puede, en algunos casos, ser muy irregular, por lo que coordenadas promedio, podrían no ubicarse en el centro de la provincia.

donde ρ y ρ_i representan los respectivos coeficientes. En este caso, ρ se encuentra definido de tal manera que: $\rho = \frac{\sum_{i=1}^n (y_i - \bar{y})(y_{i+1} - \bar{y})}{\sum_{i=1}^n (y_i - \bar{y})^2}$, donde menos uno significa que existe perfecta autocorrelación espacial negativa, y uno significa que existe perfecta autocorrelación espacial positiva, y cero indica que no existe autocorrelación espacial. En el cálculo de esta matriz, se puede asignar un rango que define la contigüidad entre las provincias. En nuestro caso, optamos por un rango de 2,5 basándonos en las distancias euclidianas promedio entre provincias¹⁶.

Las matrices W y X , representan las matrices de pesos espaciales y variables explicativas respectivamente. Se puede observar que la variable dependiente también se encuentra del lado derecho de la ecuación. Lo anterior representa lo que se mencionó anteriormente, donde se quiere observar la dependencia espacial de la variable Y con respecto a su par procedente de otra provincia. Por último, se tiene ϵ que corresponde al término del error correspondiente.

Luego, esta ecuación, se estima, como se dijo anteriormente, por el método de máxima verosimilitud, lo que nos permite obtener estimadores asintóticamente eficientes y consistentes. Nuevamente se menciona que las estimaciones se realizaron con una muestra reducida. Ésta se compone de 717 observaciones, de las cuales, todas corresponden a valores positivos del salario. La reducción de la muestra se debe a las limitaciones del software utilizado. Por otro lado, la censura de la muestra, se debe a que el comando que se utiliza para realizar la regresión espacial, no permite trabajar con observaciones que no presenten valores positivos para el salario.

¹⁶ Se define distancia euclidiana como: $d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$.

SECCIÓN IV

RESULTADOS.

Los resultados que se presentan a continuación se obtuvieron utilizando datos de la encuesta CASEN 1996, CASEN 2000 y CASEN 2009. Los modelos aplicados son los que se mencionaron en la sección anterior, a los que se suma una descomposición de Blinder – Oaxaca utilizando las variables más relevantes según nuestro criterio.

4.1 Resultados para el año 1996.

4.1.1 Estimación por mínimos cuadrados ordinarios.

Los resultados de esta estimación, en la mayoría de los casos, están en línea con lo que se pensaba a priori. En ese sentido, no hay mayor sorpresa con los signos de los coeficientes y la gran mayoría de las variables resultaron ser significativas. Para empezar con el análisis, las variables correspondientes a la experiencia de cada individuo son significativas y muestran una relación positiva con el salario. Lo anterior es esperable, debido a que entre más experiencia tenga el individuo mayor será su probabilidad de obtener un salario más alto (ver tabla A1).

En el caso, de la variable dicotómica referida a la etnia, se puede ver que existe una relación negativa con el salario, y que esta variable es significativa. El coeficiente sólo explica al salario en un -5%(ver tabla A1), cabe decir que se esperaba que fuese más alto. En todo caso, habría que ver qué sucedió con los demás períodos para realizar una mejor inferencia, teniendo en cuenta que en este período la muestra es menor debido a que la construcción de la pregunta (ver en secciones anteriores). La variable correspondiente a si la persona pertenece al género masculino resultó ser altamente significativa y presenta una relación positiva con el salario. Este resultado, como se dijo a priori, significa que la variable hombre tiene una fuerte relación con el salario en nuestro país porque es sabido que a los hombres se les remunera con un mayor sueldo. Parte de esto se debe a la discriminación por género que también existe en nuestro país, y que por esos años era aún mayor.

La variable dicotómica referida a si la persona reside en un sector urbano es significativa y presenta una relación positiva con el salario. Esto se explica porque las probabilidades de encontrar un trabajo con un mejor salario crecen en lugares urbanos, ya que es en estos lugares donde podemos encontrar más empresas y oportunidades de trabajo.

A continuación, se presentan los resultados para el grupo de variables dicotómicas referidas al nivel de educación que posee cada individuo. En primer lugar, la variable más relevante de este grupo y que además resultó ser la más significativa, es la correspondiente a si la persona posee educación terciaria completa. Esta variable muestra una fuerte relación positiva con el salario. Le sigue en esta línea, la educación secundaria completa, educación secundaria incompleta y la variable de educación primaria completa. Acorde a la teoría, y lo estudiado por Mincer, entre más grados educación presenta la persona mayor debiera ser su salario. Por lo mismo es que la variable referida al nivel de educación terciaria completa es la más significativa. Si bien es cierto, se puede tener problemas de endogeneidad entre las variables de educación y salario, no se pueden descartar del todo estas variables, ya que además las usaremos para realizar la descomposición de Blinder – Oaxaca mencionada en la sección anterior.

El siguiente grupo de variables a presentar sus resultados, corresponden a dummies de provincias. Se mencionó anteriormente que, estas provincias fueron escogidas porque es en ellas donde se encuentra la mayor población residente de personas con etnia. Los resultados nos muestran que las variables Parinacota y Chiloé resultaron ser no significativas. El resto de las variables son todas significativas pero con distinto signo. Las que muestran un signo positivo con respecto al salario son las provincias de Arica y Santiago. Una explicación posible para el signo de la provincia de Arica es que ésta, al ser considerada como zona extrema, tenga salarios mayores al promedio. Sin embargo esta explicación no es concluyente y más información debiese ser recabada.

Para el caso de Santiago, el asunto parece ser más claro. Esto es debido a la gran centralización de nuestro país, problema que era más fuerte en ese período; la gran mayoría de las empresas se encontraban físicamente en Santiago (fenómeno que se mantiene en la actualidad), por lo que las posibilidades de encontrar un empleo con un salario mayor son mayores en dicha provincia que en el resto del país. Del resto de las provincias se puede destacar las de Cautín y Arauco. Ambas son significativas y muestran un signo negativo con respecto al salario. La intención de destacar estas provincias es porque se encuentra la mayor población de mapuches residentes del país. Por lo tanto, si existiese discriminación étnica, debiera ser más fuerte en estas provincias.

Por último, el siguiente grupo de variables dicotómicas, es aquel correspondiente a la rama o actividad económica en la cual participa cada individuo. Este grupo se refiere a las principales actividades económicas realizadas en nuestro país. Los resultados arrojaron que las siguientes variables no son significativas: sector industrial, servicios básicos y comercio. Las demás variables

resultaron ser significativas con distintos grados de significancia. La variable agricultura muestra una relación negativa con el salario, lo cual era de esperarse debido a las remuneraciones bajas que normalmente se obtienen en este sector. Por el contrario, la variable minería muestra una relación positiva con el salario. Lo anterior se debe a que los salarios son normalmente más altos en este sector debido a los grandes riesgos del trabajo diario en las minas. El sector comercio muestra una relación positiva con el salario, lo anterior se explica debido al gran crecimiento que estaba teniendo el país durante ese período, lo que repercutió en un repunte de la industria del comercio en esos años. Otras variables que muestran una relación positiva con el salario son el transporte y los servicios financieros. Para este último, el resultado está dentro de lo esperado porque para este sector se requiere gente con mayores estudios, pero para el sector transporte no se tiene una explicación clara del signo obtenido. Por último, para el sector servicios, se observa una relación negativa con el salario. Para esta variable tampoco se tiene una explicación clara de su resultado.

4.1.2 Estimación de sesgo de selección por máxima verosimilitud.

En primer lugar se comentan los resultados de la ecuación de participación estimada. Podemos observar que la variable dicotómica etnia es altamente significativa y disminuye la probabilidad de participación de las personas en un 65,38% (ver tabla 4), lo que es esperable, ya que durante el período estudiado no existían leyes que apuntaran a mejorar el trato hacia las personas con origen étnico.

Con respecto a la variable, si la persona reside en un sector urbano o no, se observa como la más significativa de todas las variables, y explica positivamente la probabilidad de participación. Esto está en línea con lo esperado, ya que si una persona reside en un sector urbano tiene más probabilidades de participar en el mercado laboral debido a que es en estos sectores donde existe más trabajo. Le sigue a esta variable, la dummy correspondiente a si la persona es hombre, que afecta positivamente en la decisión de trabajar de las personas, esto puede ser explicado, porque en este período no existían mayores beneficios para la mujer en el mundo laboral. En lo referente al número de personas que residen en el hogar, este resulta poco significativa. Las variables de experiencia potencial y ser jefe de hogar, resultaron ser significativas y explican positivamente la decisión de participar de las personas. Esto va en línea con la teoría, ya que se espera que el jefe de hogar contribuya en mayor medida a la mantención de éste. Por último, el número de hijos influye negativamente en la decisión de participar del mercado laboral.

Tabla 4 Estimaciones MCO y MV 1996

ln_wage	MCO	t	MV	z
exper	0,036885	46,350	0,037567	46,470
d_etnia	-0,059716	-3,540	-0,077751	-4,460
d_hombre	0,363494	44,860	0,368207	45,160
d_urbano	0,073371	7,430	0,085409	8,280
select				
ocup			3,229798	97,510
d_etnia			-0,653823	-18,770
d_urbano			0,732314	29,790
d_hombre			0,172325	6,770
numper			0,014157	1,790
exper			0,014541	6,390
jefe			0,656989	21,700
n_hijo			-0,028927	-2,670

Ahora bien, en la estimación por Máxima Verosimilitud se puede observar, que las variables referentes a la experiencia potencial del individuo, resultaron ser significativas y positivas. En cuanto, a la variable dicotómica de la etnia, resulta ser significativa y positiva (-7,8%), tomando mayores valores en la estimación de máxima verosimilitud que en la de mco (ver tabla A2).

La variable referente a si la persona corresponde a un hombre, resulta ser altamente significativa, y explica positivamente el nivel del salario. Esto está de acuerdo a lo esperado, ya que generalmente en nuestro país, los hombres tienen un mayor sueldo que las mujeres. Por último se encuentra la variable dummy correspondiente a si el individuo vive en un sector urbano, esta variable también es significativa y explica positivamente el salario.

El siguiente grupo de variables a presentar sus resultados corresponde al grupo de variables dicotómicas referentes al nivel de educación de cada individuo. Los resultados nos indican que la variable más significativa es la que corresponde a si el individuo posee educación terciaria completa. Sin embargo, el resto de las variables también resultaron ser significativa y explican de forma positiva el nivel del salario. Esto está dentro de lo esperado, ya que de acuerdo a Mincer, a más años de escolaridad mayor debiese ser el nivel del salario. La segunda variable más significativa de este grupo es la que corresponde a si el individuo tiene educación secundaria completa. Por último, le siguen a la anterior en nivel de significancia, la dummy de educación terciaria completa, educación secundaria completa y educación primaria completa.

Ahora se presentan los resultados para el grupo de variables dicotómicas correspondiente a la provincia en donde reside la persona. Dentro de este grupo, las provincias de Parinacota y Chiloé resultaron ser no significativas. El resto son significativas pero con diferente signo. Dentro de las que explican positivamente el salario encontramos a la provincias de Arica y Santiago. De la primera, se puede expresar como una explicación inicial que al ser considerada zona extrema, los salarios tienden a ser más altos dentro de estas zonas. Con respecto a Santiago, el resultado está dentro de la normalidad, ya que dentro de esta provincia existen mayores probabilidades de encontrar mejores salarios, debido a la mayor concentración de compañías y empresas dentro de esta zona. Las demás provincias resultaron con signo negativo, es decir explican negativamente el salario, dentro de estas provincias, destacan las provincias de Cautín y Malleco, ambas son altamente significativas, pero lo más importante es que ambas concentran la mayor población de mapuches residentes del país. Lo anterior, nos muestra la importancia de el estudio que estamos realizando ya que se puede observar que los salarios en promedio son mucho más bajos que en otras zonas.

Por último se presentan los resultados para el grupo de variables dummy referentes al área o sector productivo en donde se desempeña la persona. Dentro de este grupo resultaron ser no significativas las dummy correspondientes al sector industrial, sector de servicios básicos y construcción. El resto de las variables resultaron ser todas significativas con distintos grados. Con respecto a la rama de la agricultura, ésta explica negativamente al salario, lo cuál puede ser debido a que en general este sector presenta remuneraciones más bajas que otros sectores debido a la calidad del trabajo que se realiza.

Por otro lado, el sector de la minería presenta una relación positiva con el salario. Lo anterior es esperable debido a los altos salarios que suelen pagar en este sector, debido a las condiciones de trabajo. El sector del Comercio también explica de forma positiva al salario. Esto se podría explicar debido al gran momento que estaba pasando la economía de nuestro país (previo a la crisis asiática) durante el período estudiado, recordando que el comercio en líneas generales va en sintonía con el momento por el cual está cursando la economía. Además, tanto el sector transporte como el de servicios financieros, explican positivamente el salario. Sin embargo, no encontramos una explicación plausible para el primer sector. En el caso del segundo, se podrían esperar mayores salarios debido a que se necesita gente más capacitada para trabajar en este

sector. Por último, el sector de Servicios muestra una relación positiva con el salario, sin embargo no se puede argumentar el porqué de este resultado de forma categórica.

4.2 Resultados para el año 2000.

4.2.1 Estimación por mínimos cuadrados ordinarios.

Podemos observar que el modelo en su totalidad es significativo y que el R-cuadrado resultó ser de 0,3474 (ver tabla A3). Para comenzar el análisis de este período, tenemos en primer lugar a las variables correspondientes a la experiencia potencial de cada individuo. Este par de variables son significativas, e incluso resaltan dentro del grupo de las más significativas (teniendo en cuenta el valor t de cada una de estas). Además, los signos de éstas son positivos en ambos casos. Lo anterior está dentro de lo esperado, ya que a más experiencia tenga el individuo, mayor debería ser el salario de este.

Continuando con el análisis, se presentan tres variables dicotómicas correspondientes a características personales del individuo. En primer lugar tenemos una dummy que controla por el efecto de la descendencia étnica, es decir, el parámetro refleja si el individuo posee descendencia étnica o no. El resultado indica, que la descendencia étnica afecta negativamente el salario en un 8,8% (ver tabla A3), además, se muestra que esta variable es significativa. Lo anterior indica, que podríamos estar en presencia de discriminación étnica. En segundo lugar, se presenta la variable dicotómica referente a si el individuo pertenece o no al género masculino. La estimación nos muestra que ser hombre afecta positivamente el salario (ver tabla A3).

Por último, se tiene la variable dicotómica correspondiente a si la persona vive en un sector urbano o no. Esta variable también resultó ser significativa, su parámetro es positivo y explica el nivel del salario en un 7,5%.

Con respecto a las dummy de escolaridad; la educación primaria completa es significativa y afectando positivamente el salario. En el caso de la educación secundaria, la variable de secundaria incompleta y completa son significativas, y afectan positivamente el salario. En el tercer nivel, tenemos una dummy para observar si la persona posee educación terciaria incompleta y otra dummy para ver si posee educación terciaria completa. Nuevamente, ambas variables son significativas, y en ambos casos la relación con el salario es positiva.

El siguiente grupo de variables dicotómicas, es el que corresponde a las provincias escogidas para este estudio. Como se mencionó anteriormente, se escogieron ciertas provincias del norte y del

sur del país, porque son las que presentan un mayor número de población indígena residente. Dentro de las provincias escogidas para el norte, sólo la provincia de Arica resultó ser no significativa, además presenta una relación negativa con el salario. A diferencia de lo anterior, la provincia de Parinacota resultó ser significativa y además presenta una relación positiva con el salario. Lo anterior se podría deber a que esa zona es considerada como zona extrema, por lo que los salarios son mayores que el promedio, y también se debería tener en cuenta el efecto de las mineras dentro de esa zona (las cuales se caracterizan por otorgar mayores sueldos a sus trabajadores).

Dentro de las provincias del sur, la mayoría son significativas, salvo por las provincias de Osorno y Valdivia. Además los resultados arrojaron que todas éstas provincias muestran una relación negativa con el salario salvo por Santiago, que resultó ser la más significativa y la única que muestra una relación positiva con el salario. Lo anterior, podría deberse a la gran centralización del mercado laboral que existe en nuestro país, ya que en Santiago se encuentran la mayoría de las compañías y empresas del país. Por lo anterior es que aumenta la probabilidad de encontrar trabajo a un mejor salario. Otras provincias que son altamente significativas en comparación con el resto, son las provincias de Cautín y Malleco. Esto nos llama la atención ya que es en estas provincias donde se encuentra la mayor población de mapuches residentes, por lo que el efecto de la etnia se podría ver reforzado en esta zona.

Por último, el siguiente grupo de variables binarias es el que hace referencia a la rama o actividad económica en la que participa el individuo. Dentro de este grupo, podemos observar que las variables correspondientes al sector industrial, servicios básicos, construcción, comercio y transporte resultaron ser no significativas. El resto de las variables son todas significativas pero con diferente resultado. La variable correspondiente a la agricultura muestra una relación negativa en alrededor de un 20% con respecto al salario (ver tabla A3). Lo anterior se debe a que en este sector los salarios siempre han sido bajos para las personas que poseen un nivel de educación bajo. Vale la pena recalcar que, es en este sector, se puede encontrar la mayor cantidad de personas trabajando con descendencia étnica.

Por otro lado, tenemos la variable referente a la minería que tiene una relación positiva con el salario en alrededor de un 22%. A diferencia de la agricultura, en la minería los salarios son mayores debido a las exigencias y riesgos de este sector. Otra variable que muestra una relación positiva con el salario es la correspondiente al sector servicios financieros. En este sector, los

salarios también suelen ser más altos que el promedio, ya que se requiere de un nivel educacional alto para participar de él. La última variable se refiere al trabajo en servicios. Esta muestra una relación negativa con el salario, pero la razón de este resultado no es clara ya que es un sector bastante amplio. En definitiva, el valor de cada actividad económica difiere entre cada una, debido a las características particulares de cada sector y a las necesidades y requerimientos educacionales que éstos tienen.

4.2.2 Estimación sesgo de selección por máxima verosimilitud.

Los resultados arrojados por esta metodología son similares a los que se obtienen por MCO, en cuanto a que éstos representan una estimación más exacta, ya que se estima los parámetros de interés en conjunto con una ecuación de participación, la cual controla por si el individuo participa o no del mercado laboral.

Tabla 5 Estimaciones MCO y MV 2000

In_wage	MCO	t	MV	z
exper	0,033206	59,160	0,034240	59,920
d_etnia	-0,088447	-8,220	-0,107592	-9,830
d_hombre	0,343324	58,040	0,341738	57,630
d_urbano	0,074574	12,130	0,098892	14,940
select				
ocup			3,004522	98,050
d_etnia			-0,412957	-18,330
d_urbano			0,685433	30,890
d_hombre			-0,047322	7,170
numper			0,014445	1,800
exper			0,015073	6,520
jefe			0,405291	21,780
n_hijo			-0,034021	-2,790

Continuando con esto, los resultados de la ecuación de participación nos muestra, que la variable dicotómica referente a si la persona reside o no en un sector urbano es la más significativa, explicando alrededor de un 69% (ver tabla 5) el hecho de que la persona se encuentre participando del mercado laboral. Por otro lado, la variable que resultó ser menos significativa es la variable binaria referida a si la persona reside en la provincia de Arauco. Aunque no es una variable que se pueda descartar del todo, el resultado de ésta nos dice que presenta una relación positiva con la decisión de participar del mercado laboral.

El resto de las variables son todas significativas. En primer lugar, tenemos a las variables relacionadas con la experiencia potencial de las personas. Éstas están relacionadas positivamente con la decisión de estar en el mercado laboral, lo que está en línea con lo que se esperaba, ya que a mayor experiencia posea la persona, aumenta la probabilidad de que ésta se encuentre participando en el mercado del trabajo. Una variable que nos llamó bastante la atención es la dummy correspondiente a la etnia. Esta variable es altamente significativa, y muestra una fuerte relación negativa con la decisión de participar del mercado laboral (ver tabla 5). Lo anterior, nos da un indicio de que la persona con ascendencia étnica, podría auto-discriminarse a priori debido a su condición de origen.

La siguiente variable, es la binaria correspondiente a si la persona pertenece al género masculino. Para nuestra sorpresa, esta variable presenta una relación negativa con la decisión de participar, lo que va en contra de lo que esperábamos, sin embargo, no muestra una relación fuerte con la variable dependiente y su grado de significancia es menor al esperado. Por otro lado, presentamos la variable correspondiente al número de personas que viven en el hogar. Esta variable, es significativa, pero no muestra una relación fuerte con la variable dependiente, a pesar de que influye positivamente en dicha variable. Sin embargo, no parece ser una variable que sea determinante en la decisión de participar en el mercado laboral, al menos no se muestra así en nuestros resultados.

La variable binaria correspondiente a si la persona es jefe de hogar es altamente significativa, y muestra una fuerte relación positiva con la decisión de participar. Esto es como lo habíamos pensando, ya que lo normal sería que la persona considerada como jefe de hogar participe del mercado laboral, ya que ésta es generalmente la encargada de mantener a su familia. La variable número de hijos, es significativa, y muestra una relación negativa con la decisión de participar, pero no se observa que ésta sea una variable determinante de esta decisión. Esto difiere en algo a lo que se pensaba a priori ya que esperábamos que la cantidad de hijos en el hogar sea una variable de peso a la hora de decidir si participar o no del mercado laboral. Por último, la variable referente a la provincia de Cautín es significativa, y muestra una relación negativa con la decisión de participar, lo que está en línea con lo esperado, ya que es una zona más bien rural donde las personas realizan actividades propias en vez de participar del mercado laboral.

En lo siguiente presentaremos los resultados de la estimación por máxima verosimilitud. Dentro de los resultados, podemos observar que los patrones y tendencias son similares a los mostrados

por la estimación MCO. La variable dummy referida a si la persona posee un nivel de educación terciaria resultó ser la más significativa de todas. Por otro lado, la variable dummy correspondiente al sector de servicios básicos resultó ser la menos significativa, con un valor bastante pobre.

Nuevamente, se presentan las variables de experiencia potencial. Éstas, al igual que al estimar por MCO, son significativas, y muestran una relación positiva con el salario. Lo anterior se explica de la misma manera mencionada en la estimación de MCO (ver tabla 5).

La variable referente a la etnia es significativa y muestra una relación negativa con respecto al salario, donde éste se ve explicado en un 10,8% por la variable explicativa mencionada (ver tabla A4). Esto es consecuente con la estimación por MCO, y con el período anterior, eso sí, se observa un mayor coeficiente para este período. La variable correspondiente a si la persona es de sexo masculino es significativa, y está relacionada positivamente con el salario. Por otra parte, la variable referente a si la persona vive en un sector urbano también es significativa y se muestra una relación positiva con el salario. Como se ha dicho anteriormente, es más probable que, se encuentre un empleo donde se percibe un mayor sueldo en sectores urbanos.

Continuando con los resultados, las variables dicotómicas que se refieren al nivel de educación que tiene la persona, resultan ser todas significativas. Además, todas las variables tienen una relación positiva con el salario, lo que es de esperar, ya que a mayor nivel de educación, mayor debería ser el salario de la persona. Dentro de estas variables, destaca la que se refiere a si la persona posee educación terciaria completa, ya que ésta resultó ser la más significativa. Después de ésta, le sigue la variable que se refiere a si la persona posee educación secundaria completa. Al igual que en el caso MCO, se puede observar que a la persona le va mejor si tiene educación secundaria completa y no si posee educación terciaria incompleta.

Por último, la variable que presenta un menor nivel de significancia dentro de este grupo en particular, es la variable correspondiente a si la persona posee un nivel de educación primaria completa. En definitiva, se observa la fuerte relación que existe entre el nivel de educación y el salario, tal como se puede apreciar en las ecuaciones de Mincer.

El siguiente grupo de variables, es aquel que corresponde a las provincias. Dentro de estas variables, resultaron ser no significativas las provincias de Arica, Osorno y Valdivia. El resto de las provincias son todas significativas. Dentro de este sub grupo, las provincias de Parinacota, Chiloé y

Santiago muestran una relación positiva con el ingreso. El resultado de la provincia de Santiago es consecuente con lo que se pensó a priori, donde para Santiago la relación positiva con el salario se explica por la gran concentración de compañías y empresas que existe en dicha provincia.

Por último, el siguiente grupo de variables dicotómicas es el que se refiere al sector o actividad económica en donde se desempeña la persona. Dentro de este grupo, las que resultaron ser poco y nada significativas son: el sector industrial, sector servicios básicos, construcción, comercio y transporte. Las variables agricultura y Servicios presentan una relación negativa con el salario y son ambas significativas. Las razones de este resultado son las mismas que se explicaron en la sección anterior, para el caso del sector de servicios financieros y minería, ambas variables tienen una relación positiva con el salario, y ambas son significativas. Este resultado se explica por las mismas razones que se han mencionado anteriormente.

4.3 Resultados para el año 2009.

4.3.1 Estimación por mínimos cuadrados ordinarios.

El modelo en su conjunto es significativo, y el R-cuadrado resultante es de 0.2934. Con respecto a las variables explicativas, se puede observar que, las referentes a la experiencia potencial del individuo son significativas y están en línea con lo que dicta la teoría, es decir, a mayor experiencia, mayor debería ser el salario que recibe el individuo. Respecto a la variable dicotómica correspondiente a la descendencia indígena, ésta es significativa, pero lo es en un gran menor que gran parte de las demás variables ocupadas, considerando los valores t. El resultado de esta variable, está en línea con lo predicho por nosotros, en relación a que, si la persona presenta descendencia indígena recibirá un menor sueldo en comparación a una persona no indígena. Según este modelo la disminución en el sueldo sería alrededor de un 7% (ver tabla A5). Una variable que resultó ser altamente significativa, es la dummy referente al género. Posee un valor t alto y explica alrededor de un 34% el salario recibido por el individuo. Por otro lado, la variable dummy correspondiente a si la persona vive en un sector urbano, o no también es significativa, pero su valor t es menor: sin embargo, el resultado de ésta está en línea con lo esperado, ya que se esperaría que una persona recibiera un mayor sueldo en un sector urbano.

Con respecto a las variables dicotómicas referentes al nivel de educación, todas resultaron ser significativas, pero dentro de ellas, las que más destacan son las que hacen referencia a si la persona tiene educación terciaria completa, si la persona tiene educación terciaria incompleta y a si la persona tiene educación secundaria completa. Este resultado, está acorde con lo que se

esperaba de estas variables, ya que mientras mayor educación posea el individuo, mayores serán sus probabilidades de obtener un mejor salario. Aún más, la variable dummy referente a si la persona tiene una educación terciaria completa, resultó ser la más significativa de todo el modelo. Lo anterior, resulta fácil de explicar, ya que una persona con un nivel profesional, tiene una mayor oportunidad de obtener un mejor salario que una persona que sólo posee estudios secundarios.

Por otro lado, tenemos las variables dummy correspondientes a las provincias con mayor población indígena del país. Respecto de éstas, los resultados arrojan que las provincias de Parinacota y Chiloé son poco y nada significativas. Sin embargo, el resto de las variables usadas, resultan ser significativas, donde las que más destacan, son las provincias de Cautín y Santiago. Con respecto a los resultados, todas las provincias muestran que el residir fuera de Santiago, afecta negativamente el salario que los individuos perciben, a excepción de la provincia de Arica. Sin embargo, nos llama la atención el resultado correspondiente al sector de Arica. Probablemente lo anterior se debe a un factor en particular que no estamos midiendo, o que probablemente el hecho de ser una ciudad considerada dentro del rango de “zona extrema”, hace que los salarios sean mayores (tomando en cuenta que ninguna de las demás provincias de la muestra es considerada zona extrema). Otra provincia que destaca es la de Cautín, la cual influye negativamente en el salario de un individuo en un 16% (ver tabla A5). Este resultado también nos llama la atención, ya que es dentro de esta provincia donde se encuentra la mayor cantidad de mapuches residentes. Lo anterior, puede ser un indicio de discriminación salarial por etnia, sin embargo, para que esta información sea verdadera, se requiere de un mayor estudio.

Por último, tenemos a las variables dicotómicas que hacen referencia al sector o actividad económica en donde se desenvuelven los individuos. Los resultados nos muestran que los sectores correspondientes a los servicios básicos (agua-luz-gas), a la construcción y al sector industrial, son variables no significativas. Con respecto a lo mencionado en la sección anterior, los signos correspondientes al sector financiero y minería están en línea con lo esperado, ya que en ambos casos se observa una relación positiva con respecto al salario que perciben los individuos. Incluso, la variable correspondiente al sector de la minería explica en una 35% el salario de éstos, y resultó ser la más significativa de todas. Por el contrario, la variable correspondiente al sector de la agricultura, también es significativa pero muestra una relación negativa con el salario, lo que se traduce en un valor de -16%.(ver tabla A5) Lo anterior, es importante para nosotros, porque dentro de este sector es donde se encuentra la mayor muestra de personas trabajando con

descendencia indígena; es más, los mapuches tienen una alta presencia en trabajos correspondientes a la agricultura.

4.3.2 Estimación sesgo de selección por máxima verosimilitud.

Los resultados de este modelo son algo diferentes a los entregados por MCO ya que, al introducir la ecuación de participación, nuestra medición se hace más exacta. Además, al utilizar esta estimación propuesta por Heckman, hacemos uso de los datos que aportan las variables censuradas, es decir aquellas para las cuales no se observa un ingreso salarial positivo.

Tabla 6 Estimaciones MCO y MV 2009

ln_wage	MCO	t	MV	z
exper	0,025708	54,750	0,026271	55,440
d_etnia	-0,053355	-6,620	-0,059573	-7,360
d_hombre	0,339651	67,490	0,341218	67,710
d_urbano	0,030051	5,630	0,034882	6,500
select				
ocup			3,468500	145,160
d_etnia			-0,298687	-13,980
d_urbano			0,269021	18,360
d_hombre			0,131241	8,750
numper			0,016596	2,890
exper			0,018073	13,780
jefe			0,392667	22,350
n_hijo			-0,023139	-2,740

Dentro de la ecuación de participación, podemos observar que la gran mayoría de las variables resultaron ser significativas. Dentro de este grupo, las que más destacan son la variable dicotómica correspondiente a si la persona es jefe de hogar, las variables referentes a la experiencia potencial del individuo, y las variables cualitativas correspondientes a si el individuo pertenece al género masculino, si tiene descendencia étnica y si reside en un sector urbano.

Dentro de las variables mencionadas, destaca aquella relacionada a la etnia. Esta variable, arrojó como resultado una relación negativa con la decisión de estar ocupado o no, y explica en un -28% esta decisión (ver tabla 6). Lo anterior, puede deberse a que una persona con descendencia étnica puede auto - descartarse a priori pensando en que puede ser discriminado en el mercado laboral por su condición indígena. Esto podría dar indicios de que se puede estar produciendo una propia discriminación, debido a lo que se piensa del mercado laboral. Por otro lado, las variables correspondientes al sexo y al sector urbano, están relacionadas positivamente con la decisión de

estar ocupado. Lo anterior está en línea con la teoría, ya que es en los sectores urbanos donde se encuentra la mayor probabilidad de estar empleado. En el caso del género, el resultado no sorprende en gran parte debido a la discriminación por género existente en nuestro país.

Continuando con la presentación de resultados, la estimación por máxima verosimilitud presenta diferencias sutiles con respecto a las estimaciones por mínimos cuadrados ordinarios. La variable más significativa continúa siendo la dummy correspondiente a si la persona posee estudios terciarios completos. En cambio, la variable menos significativa corresponde a la dummy que informa si la persona trabaja en el sector de servicios básicos; las demás variables arrojaron valores dentro de lo esperado. En primer lugar, las variables que miden la experiencia potencial de los individuos, son significativas, pero explican en un porcentaje muy bajo el salario de las personas.

En segundo lugar, tenemos a tres variables dicotómicas, en donde la primera es la que se refiere a la descendencia étnica de las personas. En este caso, esta variable resultó ser un poco más significativa que la resultante por MCO, y siguen teniendo el mismo signo aunque ahora ésta explica sólo en un -8% el valor del salario del individuo. La segunda variable dicotómica, es la que se refiere al género masculino. En este caso, la variable explica en un 34% el salario del individuo (ver tabla 6). Esta variable, también resultó ser altamente significativa, lo que en teoría era de esperarse, potencialmente debido a la existencia de discriminación por género existente en el país. Y en tercer lugar, se presenta la variable dicotómica que hace alusión a si la persona vive en un sector urbano. Esta variable es significativa, y logra explicar solo un 3% del ingreso salarial (ver tabla A6). Como se ha explicado antes, lo normal es que esta variable tenga una relación positiva con el ingreso, ya que es en sectores urbanos donde se pueden encontrar más empleos con mejores salarios.

Las variables dicotómicas que intentan explicar el efecto de tener un cierto nivel de educación en el salario son todas significativas, donde la que más destaca es la que se refiere a si el individuo posee educación terciaria. Le siguen a esta variable, las dummy correspondientes a si el individuo posee educación terciaria incompleta, y si el individuo posee educación secundaria completa. Como es de esperar, un individuo con educación terciaria completa tiene mayores probabilidades de obtener un salario mayor. En este caso, la variable correspondiente, tiene el valor t más alto de la muestra. En consecuencia, individuos con un menor nivel educacional, tienen menores probabilidades de conseguir un buen salario. Esto queda reflejado en los resultados obtenidos, ya

que las variables que reportan un nivel educacional menor, son menos significativas, y explican en un menor porcentaje el salario de cada individuo.

Por el lado de las variables dummy correspondientes a las provincias, no observamos diferencias sustantivas con respecto a MCO. Parinacota y Chiloé muestran ser poco significativas. Sin embargo, el resto de ellas sí son significativas, en donde nuevamente las que destacan son las provincias de Cautín y Santiago. La última, junto con la provincia de Arica son las únicas que muestran una relación positiva con el ingreso. Esto, como se explicó anteriormente, se debe efecto del grado de centralización de nuestro país (para el caso de Santiago), ya que es en la capital donde se encuentra el mayor número de empresas, y en consecuencia se tienen las mayores probabilidades de encontrar un empleo a un mejor salario. En el caso de Arica, una posible explicación a este resultado, es el hecho de que ésta provincia es considerada como zona extrema, en donde los salarios son mayores incluso hasta en un 40% (para el caso del sector público).

Lo que nos llama nuevamente la atención, es que las provincias del sur, particularmente, Cautín y Malleco, obtienen el peor desempeño. En estas provincias, el ingreso se ve explicado en un -16% y -17%, respectivamente. Nuestro interés se focaliza en que según la información proporcionada por la encuesta, es justamente en esas provincias, donde podemos encontrar la mayor población de mapuches residentes, por lo que podríamos encontrar indicios de discriminación étnica y además de una cierta correlación entre estas comunas (ver tabla A6).

Por último, se entregan los resultados acerca de las variables dicotómicas que hacen referencia al sector o actividad económica en la que se desempeña el individuo. Los resultados arrojan que las variables dummy del sector de servicios básicos, sector industrial, construcción y comercio no son significativas. Por otro lado, las variables correspondientes al sector agrícola, minería, transporte, sector financiero y servicios son todas significativas. Dentro de este grupo, las que presentan una relación positiva con el ingreso son el sector financiero, el transporte y la minería. Con respecto a la minería, es de esperar que se obtengan mayores salarios, puesto que se trabaja en condiciones adversas y responde a los grandes ingresos que están obteniendo las empresas mineras por concepto de alza en los precios de los metales, particularmente el cobre y el molibdeno. Por el lado, del sector financiero, no sorprende este resultado, los ingresos que se obtienen en este sector son mayores que en otros, entre otra cosas porque se requiere de un mayor nivel

educacional para participar en el área financiera, por lo que esto debe estar correlacionado con el nivel educacional del individuo, lo que también responde a su vez, a un aumento del salario.

4.4 Resultados descomposición de Blinder – Oaxaca.

Recordemos que este método nos muestra la existencia de un diferencial en una variable a elección, en este caso el salario de los individuos. Además nos dice cuánto de este diferencial se ve explicado por las variables que elegimos para estimar, es decir, qué porcentaje se explica por diferencias en capital humano y por diferencias en el retorno al capital humano. Con respecto a esto último, en general, la literatura asigna las diferencias en el retorno al capital humano a la existencia de discriminación salarial.

Tabla 7 Descomposición Blinder-Oaxaca

Blinder-Oaxaca decomposition			
1: d_etnia = 0			
2: d_etnia = 1	1996	2000	2009
Difference	38,07	35,01	24,02
Descomposición			
Explicado	60,21	48,61	58,09
No explicado	39,79	51,39	41,91

En el año 1996, se observa que existe una diferencia del 38,07% entre el grupo de los que no poseen origen étnico y los que sí la poseen (ver tabla 7). De este porcentaje, un 60,21% se ve explicado por las diferencias en el capital humano y el 39,79% restante no se puede atribuir a diferencias en capital (ver tabla 7). Como dijimos anteriormente, esto corresponde a la existencia de discriminación salarial por etnia.

En el caso del año 2000, el resultado de esta descomposición nos muestra una diferencia importante de alrededor de un 35,01% entre las personas con descendencia étnica y sin descendencia étnica, en donde las primeras, son las que perciben un menor salario en comparación con el segundo grupo. Se puede apreciar que, un 48,01% del diferencial se ve explicado por diferencias en el capital humano, es decir, diferencias en el nivel del conocimiento entre cada grupo. Mientras, un 51,39% restante no pudo ser explicado por las variables, más bien se le adjudica este porcentaje a las diferencias en los retornos al capital humano, en este caso, discriminación étnica. Por lo tanto, durante este período, si se hubiesen mejorado las condiciones educacionales para las personas con descendencia étnica, gran parte de la brecha mostrada por la descomposición se vería reducida (ver tabla A7).

Por último, para el 2009, la descomposición muestra que existe una diferencia en total de un 24,02% entre los que no poseen descendencia étnica y los que sí la poseen. De esta diferencia, un 58,09% logró ser explicado por las diferencias en el capital humano, es decir por las diferencias en las variables que nosotros estimamos eran las más significativas. En cambio, el 41,91% restante es explicado por las diferencias al retorno del capital humano, es decir, se debe atribuye a discriminación étnica (ver tabla A7).

4.5 Resultados de los test autocorrelación espacial 2009

Luego de realizar la metodología de máxima verosimilitud para el año 2009, analizamos la posibilidad de existencia de autocorrelación espacial entre las distintas provincias, con el objetivo de mejorar la estimación. Para poder testear la presencia de este tipo de autocorrelación, utilizamos los estadísticos que sugiere la literatura, que corresponden a la I de Moran y C de Geary.

Previo a esto, se debió construir la matriz de pesos espaciales, que tal como señalamos en la metodología, nos permite determinar cuáles provincias son contiguas y cuáles no. El test se realiza sobre la variable dependiente (ln_wage), ya que lo que se quiere evaluar es la posibilidad de que el salario de una provincia se vea influenciado por el de otras provincias contiguas. Si bien esto podría haberse realizado con comunas, la no existencia de una encuesta CASEN georeferenciada, limita mucho estas evaluaciones. El primer test que presentaremos corresponde a la I de Moran:

Tabla 8 I de Moran 2009

Moran's I					
Variables	I	E(I)	sd(I)	z	p-value*
ln_wage	-0,017	0,001	0,002	7,752	0,000

*1-tail test

Como podemos ver, según el test de Moran, existe evidencia de que hay autocorrelación negativa en el salario, ya que valores entre -1 y 1 reflejan la presencia de autocorrelación, lo que nos indica que nuestra estimación podría mejorarse incorporando la matriz de pesos espaciales al modelo. Para esto, se realizó la estimación de un modelo SAR (spatial autocorrelation regression). Para verificar nuestros resultados, utilizaremos también el test de Geary, que mostramos a continuación:

Tabla 9 C de Geary 2009

Geary's c					
Variables	c	E(c)	sd(c)	z	p-value*
ln_wage	0,928	1,000	0,042	-1,700	60,044

*1-tail test

Según el estadístico C de Geary, podemos corroborar la presencia de autocorrelación negativa. Aún así, es importante destacar, que el test de Moran es más sensible a la autocorrelación espacial global, que el test de Geary¹⁷.

4.3.5 Resultados para la estimación de autocorrelación espacial (SAR) 2009

Tabla 10 Estimación SAR y MV 2009

ln_wage	SAR*	z	MV*	z
exper	0,025812	4,88	0,025680	4,84
d_etnia	-0,166344	-1,78	-0,136039	-1,47
d_hombre	0,287518	5,18	0,285964	5,13
d_urbano	0,152602	2,47	0,154396	2,49

*con muestra reducida 2009

Como se mencionó anteriormente, la motivación de realizar una estimación corrigiendo por autocorrelación espacial, es la de mejorar nuestros resultados. Sin embargo, éstos no pudieron mejorarse, entre otras cosas, debido a la disminución del número de observaciones utilizadas (cabe recordar que las observaciones utilizadas sólo presentan valores positivos para el salario). Lo anterior se debe a que el software utilizado no nos permitió utilizar más variables que las comentadas en la sección anterior. Por lo tanto, realizamos una estimación de máxima verosimilitud con la misma cantidad de datos utilizados en la regresión espacial (es decir, 717 observaciones), para hacer los resultados comparables (ver tabla A10, Anexos).

En primer lugar, se puede decir que, para la regresión espacial, tanto las variables de educación como las dummy correspondientes al sexo, etnia y urbanidad, son significativas, pero con diferentes niveles de significancia (ver tabla A9, Anexos). Sin embargo, para el caso de la estimación del por máxima verosimilitud, sólo variable referente a la etnia resultó ser no significativa, las demás variables presentadas son todas significativas (ver tabla A9, Anexos). Es

¹⁷ Ver sección 3.4

decir, en una muestra reducida como la utilizada en este caso, se puede ver que el coeficiente correspondiente a la etnia es más alto en la estimación espacial que en la otra. Aun más, las demás variables también presentan coeficientes mayores en la primera estimación que en la segunda. Sin embargo, los signos de los coeficientes son los mismos en cada estimación. Por último, en los grupos de variables dicotómicas referentes a la provincia de procedencia y a la actividad económica en la que se desenvuelve, para ambos casos, los valores resultaron ser poco significativos (ver tablas en la sección anexo).

Por otro lado, un resultado bastante importante para nosotros a observar, es el del coeficiente rho. Se puede observar que valor de rho es significativo, y que refleja una mínima existencia de autocorrelación espacial negativa en la muestra. La significancia de rho se puede observar con los resultados de los distintos test aplicados (entre paréntesis se presentan los p-values):

Hipótesis nula: $\rho=0$

Tabla 11 Test de la estimación SAR

Wald test	chi2(1) = 6.145 (0.013)
Likelihood ratio test	chi2(1) = 6.119 (0.013)
Lagrange multiplier	chi2(1) = 6.071 (0.014)

En primera instancia, se pudo mejorar la estimación hecha por máxima verosimilitud bajo el mismo umbral de observaciones. Aún cuando, la cantidad de observaciones fue mínima, los resultados nos dejan bastante tranquilos, ya que se logró dar un paso más en la mejora de las nuevas estimaciones. También, se pudo comprobar la existencia (aun que mínima) de autocorrelación espacial, lo cual también representa una mejora en nuestras estimaciones. Por lo tanto, ambos resultados representan un fuerte indicio para poder realizar mejores estimaciones en el futuro, ya que se podría decir, a priori, que no es correcto pensar que las variables de una provincia (región, etc.) no afectan a las variables de otra provincia, o que no exista correlación entre zonas colindantes de una región.

SECCIÓN V

CONCLUSIONES.

La idea del presente trabajo era la de observar la presencia de discriminación étnica en nuestro país para distintos períodos de tiempo. Para ello, utilizamos la encuesta CASEN de los años 1996, 2000 y 2009. Nuestros resultados mostraron que existe discriminación por etnia en nuestro país. Sin embargo, ésta ha ido cambiando en el tiempo. Si bien para el último período utilizado, el año 2009, se puede observar una menor brecha salarial entre las personas con y sin etnia, la presencia de esta brecha nos muestra algo que no es bueno para el crecimiento de nuestra sociedad. Se puede verificar que la diferencia se ve sólo explicada, en promedio, en un 50% por las variables explicativas, por lo que el resto se puede atribuir a la existencia de discriminación. Lo anterior, nos dice que, con sólo mejorar ciertas variables de interés, como la educación para los grupos étnicos, sólo se reducirá la brecha a la mitad. Por ende, se necesitará más que mejorar la educación, por ejemplo, para poder disminuir la discriminación salarial étnica que existe en nuestro país.

A través de nuestras estimaciones, pudimos ver que el pertenecer a una etnia, empeora las posibilidades de acceder a un mejor salario, aún más, reduce las posibilidades de poder participar del mercado laboral. La idea de utilizar los tres períodos mencionados anteriormente, es poder comparar y cuantificar si la discriminación ha ido disminuyendo con el tiempo. Con la creación de ciertas entidades como la CONADI, y la puesta en práctica de nuevas leyes¹⁸ a favor de las etnias de nuestro país (la mayoría de éstas fueron creadas después del año 2000), hubiésemos pensando que el nivel de discriminación se vería reducido entre los años 1996 y 2009. Sin embargo, esto no ha pasado. Se puede verificar que la brecha se ha ido reduciendo, pero la existencia de la discriminación se mantiene fuerte.

Por otro lado, los resultados de nuestro análisis espacial, muestran una mínima existencia de autocorrelación espacial. Nuestro objetivo, al realizar este tipo de estimación, era comprobar la existencia de este tipo de autocorrelación. Por lo tanto, podemos decir que las estimaciones normales, es decir las que realizamos por mínimos cuadrados ordinarios y por máxima verosimilitud, no están del todo bien. En nuestra regresión espacial, el salario se ve influenciado por el salario de otras provincias (las más cercanas). Si bien el coeficiente obtenido es pequeño, debido a que contamos con menos de mil observaciones para estimar, es un piso para poder continuar con más estudios por el lado de esta rama de la econometría.

¹⁸ <http://www.conadi.gob.cl/>

Proponemos para estudios posteriores, continuar investigando la discriminación étnica utilizando técnicas de econometría espacial. La idea es que estos estudios se hagan a un nivel espacial más específico, por ejemplo, entre comunas donde exista mayor concentración de población indígena. Sumado a lo anterior, se podría hacer un estudio del Gran Santiago, la gracia de esto es que se tienen varias comunas que colindan entre sí, por lo que la presencia de autocorrelación espacial debería ser más fuerte que la que obtuvimos en nuestro trabajo. Aún así, los resultados que, potencialmente se puede obtener en esta provincia, se verán afectados por la baja presencia de personas con etnia.

Finalmente, el tema de la discriminación étnica, como toda discriminación, requiere más que sólo mejorar la educación para estos grupos. Se necesita una inclusión profunda de los pueblos étnicos en nuestra sociedad; se necesita aumentar las oportunidades para ellos. La brecha sigue manteniéndose alta, aunque ha ido disminuyendo en el tiempo, no ha mostrado una fuerte reducción en los últimos 15 años. La idea es que las políticas sociales apunten a mejorar este tema, quizás aumentando la integración de éstos a la sociedad más moderna, sin dejar de lado sus raíces.

Chile, en la última encuesta CASEN (2009), presenta una población indígena mayor al 10%. Esta cifra no es despreciable, si se tiene una población de más de un millón y medio de personas que presenta un origen étnico, por lo que la base para generar nuevos estudios es bastante amplia. Dejamos propuesto, entonces, continuar con la investigación en este tema, además de ahondar en los estudios de econometría espacial. La necesidad de mayores trabajos en discriminación étnica en Chile es grande, ya que con ellos se podrá orientar de mejor manera las políticas públicas para el futuro. Existen los datos y las herramientas, sólo falta la motivación para abordar este tema.

SECCIÓN VII

REFERENCIAS.

ATAL, JUAN PABLO, ÑOPO, HUGO AND WINDER, NATALIA, “New Century, Old Disparities: Gender and Ethnic Wage Gaps in Latin America”, Inter American Development Bank, Department of Research and Chief Economist , 2009.

BALAKRISHNAN N., SCHUCANY WILLIAM, “Introduction to Spatial Econometrics”, Taylor and Francis Group, LLC, 2009.

BENAVENTE, JOSE MIGUEL, “Microeconometría Aplicada Notas de Clases”, Departamento de Economía, Universidad de Chile, 8 de marzo de 2010.

BERTRAND MARIANNE AND MULLAINATHAN SENDHIL, “Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination, June 20, 2004.

CAMERON, A. COLIN AND TRIVEDI, PARVIN K., “Microeconometrics Using Stata”, A Stata Press Publication, Stata Corp LP, College Station, Texas.

JANN BEN, “A Stata implementation of the Blinder – Oaxaca decomposition”, ETH Zurich sociology working paper No. 5, may2008.

JARA, BENJAMIN A., “Discriminación en mercados laborales por genero y etnia en Chile”, RIMISP, Centro Latinoamericano para el Desarrollo Rural, 2010.

LeSAGE, JAMES P., “Spatial Econometrics”, Department of Economics, University of Toledo, december 2008.

P. DE JONG, C. SPRENGER, AND F VAN VEEN, “On extreme values of Moran’s I and Geary’s C”.

PAZ MORAGA, CAROLINA Y MIZALA SALCES ALEJANDRA, “Discriminación Salarial entre la Población Indígena Mapuche y la no Indígena en Chile”, tesis para optar al grado de Magister en economía aplicada, mayo 2008.

REED DEBORAH AND CHENG JENNIFER, “Racial and Ethnic Wage Gaps in the California Labor Market”, Public Policy Institute of California, 2003.

VÁSQUEZ NÚÑEZ, JAVIERA, "Estadística y Econometría Aplicada usando Stata", Biblioteca del Congreso Nacional.

ANEXOS.

Tabla A1 Ecuación del Salario 1996, MCO

Source	SS	df	MS	Number of obs =	42971
Model	11.432,8757	28	408,3170	F(28, 43538) =	847
Residual	20.711,5012	42942	0,4823	Prob > F =	0,0000
Total	32.144,3769	42970	0,7481	R-squared =	0,3557
				Adj R-squared =	0,3553
				Root MSE =	0,6945

In_wage	Coef.	Std. Err.	t	P>t
exper	0,036885	0,000796	46,350	0,000
exper2	-0,000419	0,000014	-29,660	0,000
d_etnia	-0,059716	0,016878	-3,540	0,000
d_hombre	0,363494	0,008103	44,860	0,000
d_urbano	0,073371	0,009881	7,430	0,000
primc	0,174166	0,012309	14,150	0,000
seci	0,400550	0,010774	37,180	0,000
secc	0,630386	0,011230	56,130	0,000
teri	1,006254	0,019295	52,150	0,000
terc	1,424371	0,013740	103,660	0,000
ari	0,254779	0,022032	11,560	0,000
pari	0,001260	0,049889	0,030	0,980
arau	-0,111278	0,023489	-4,740	0,000
cau	-0,127443	0,014795	-8,610	0,000
malle	-0,201385	0,022733	-8,860	0,000
chil	0,023860	0,038146	0,630	0,532
osor	-0,180781	0,029409	-6,150	0,000
stgo	0,255968	0,008849	28,930	0,000
vald	-0,071360	0,026693	-2,670	0,008
agri	-0,109403	0,046939	-2,330	0,020
minas	0,310165	0,050516	6,140	0,000
industria	0,010896	0,046853	0,230	0,816
elec	0,053059	0,061485	0,860	0,388
construc	0,083965	0,047326	1,770	0,076
comer	0,109714	0,046649	2,350	0,019
trans	0,158589	0,047653	3,330	0,001
fin	0,274229	0,048702	5,630	0,000
serv	-0,104577	0,046514	-2,250	0,025
_cons	10,306750	0,048439	212,780	0,000

Tabla A2 Ecuación del Salario 1996, Sesgo de Selección por Máxima Verosimilitud

Heckman selection model	Number of obs	=	48694
(regression model with sample selection)	Censored obs	=	5425
	Uncensored obs	=	43269
	Wald chi2(28)	=	22770,1
Log likelihood = -53938,39	Prob > chi2	=	0

In_wage	Coef.	Std. Err.	Z	P>z
Exper	0,037567	0,000808	46,470	0,000
exper2	-0,000433	0,000014	-29,940	0,000
d_etnia	-0,077751	0,017423	-4,460	0,000
d_hombre	0,368207	0,008153	45,160	0,000
d_urbano	0,085409	0,010314	8,280	0,000
Primc	0,170784	0,012407	13,770	0,000
Seci	0,402412	0,010835	37,140	0,000
Secc	0,631551	0,011321	55,780	0,000
Teri	1,011625	0,019372	52,220	0,000
Terc	1,431917	0,013828	103,550	0,000
Ari	0,255009	0,022078	11,550	0,000
Pari	-0,000858	0,050001	-0,020	0,986
Arau	-0,110013	0,023575	-4,670	0,000
Cau	-0,140092	0,015136	-9,260	0,000
Malle	-0,200072	0,022837	-8,760	0,000
Chil	0,023689	0,038295	0,620	0,536
Osor	-0,183047	0,029626	-6,180	0,000
Stgo	0,250893	0,008918	28,130	0,000
Vald	-0,070831	0,026722	-2,650	0,008
Agri	-0,109858	0,047003	-2,340	0,019
minas	0,308181	0,050588	6,090	0,000
industria	0,011006	0,046915	0,230	0,815
Elec	0,052180	0,061572	0,850	0,397
construc	0,084112	0,047390	1,770	0,076
comer	0,109356	0,046710	2,340	0,019
Trans	0,157832	0,047717	3,310	0,001
Fin	0,273021	0,048769	5,600	0,000
Serv	-0,117234	0,046613	-2,520	0,012
_cons	10,282600	0,048716	211,070	0,000
select				
ocup	3,229798	0,033121	97,510	0,000
d_etnia	-0,653823	0,034836	-18,770	0,000
d_urbano	0,732314	0,024581	29,790	0,000
d_hombre	0,172325	0,025460	6,770	0,000

numper	0,014157	0,007929	1,790	0,074
exper	0,014541	0,002274	6,390	0,000
exper2	-0,000541	0,000036	-15,010	0,000
Arau	0,137175	0,075949	1,810	0,071
Cau	-0,600472	0,033146	-18,120	0,000
Jefe	0,656989	0,030280	21,700	0,000
n_hijo	-0,028927	0,010833	-2,670	0,008
_cons	-2,107327	0,055066	-38,270	0,000
<hr/>				
/athrho	0,132312	0,029564	4,480	0,000
/lnsigma	-0,362248	0,003459	-104,730	0,000
<hr/>				
Rho	0,131546	0,029053		
Sigma	0,696109	0,002408		
lambda	0,091570	0,020272		

Tabla A3 Ecuacion del Salario 2000, MCO

Source	SS	df	MS	Number of obs =	76028,00
Model	17.947,6754	28	640,9884	F(28, 43538) =	1444,95
Residual	33.713,7045	75999	0,4436	Prob > F =	0,0000
Total	51.661,3799	76027	0,6795	R-squared =	0,3474
				Adj R-squared =	0,3472
				Root MSE =	0,6660

ln_wage	Coef.	Std. Err.	t	P>t
exper	0,033206	0,000561	59,160	0,000
exper2	-0,000413	0,000009	-44,650	0,000
d_etnia	-0,088447	0,010761	-8,220	0,000
d_hombre	0,343324	0,005916	58,040	0,000
d_urbano	0,074574	0,006146	12,130	0,000
primc	0,205217	0,007854	26,130	0,000
seci	0,356252	0,008115	43,900	0,000
secc	0,653604	0,007915	82,570	0,000
teri	0,978266	0,015117	64,710	0,000
terc	1,553823	0,010812	143,710	0,000
ari	-0,038529	0,028447	-1,350	0,176
pari	0,105835	0,044656	2,370	0,018
arau	-0,132844	0,017428	-7,620	0,000
cau	-0,146712	0,011437	-12,830	0,000
malle	-0,191279	0,015140	-12,630	0,000
chil	0,055713	0,015917	3,500	0,000
osor	-0,012099	0,019080	-0,630	0,526
stgo	0,155130	0,007338	21,140	0,000
vald	-0,011127	0,017186	-0,650	0,517
agri	-0,204059	0,045056	-4,530	0,000
minas	0,227932	0,047764	4,770	0,000
industria	-0,071216	0,045241	-1,570	0,115
elec	-0,006661	0,052695	-0,130	0,899
construc	-0,024288	0,045513	-0,530	0,594
comer	0,008769	0,045055	0,190	0,846
trans	0,047585	0,045706	1,040	0,298
fin	0,126957	0,046289	2,740	0,006
serv	-0,158337	0,044908	-3,530	0,000
_cons	10,669780	0,045878	232,570	0,000

Tabla A4 Ecuación del Salario 2000, Sesgo de Selección por Máxima Verosimilitud

Heckman selection model	Number of obs	=	94795
(regression model with sample selection)	Censored obs	=	18768
	Uncensored obs	=	76027
	Wald chi2(28)	=	39114,46
Log likelihood = -105412,3	Prob > chi2	=	0,00

In_wage	Coef.	Std. Err.	z	P>z
Exper	0,034240	0,000571	59,920	0,000
exper2	-0,000434	0,000009	-45,700	0,000
d_etnia	-0,107592	0,010944	-9,830	0,000
d_hombre	0,341738	0,005930	57,630	0,000
d_urbano	0,098892	0,006620	14,940	0,000
Primc	0,206182	0,007846	26,280	0,000
Seci	0,356329	0,008110	43,940	0,000
Secc	0,650181	0,007921	82,090	0,000
Teri	0,978358	0,015118	64,720	0,000
Terc	1,555675	0,010812	143,890	0,000
Ari	-0,037987	0,028398	-1,340	0,181
Pari	0,103629	0,044523	2,330	0,020
Arau	-0,130102	0,017475	-7,450	0,000
Cau	-0,154287	0,011484	-13,430	0,000
Malle	-0,191958	0,015126	-12,690	0,000
Chil	0,055229	0,015890	3,480	0,001
Osor	-0,011055	0,019058	-0,580	0,562
Stgo	0,155476	0,007342	21,180	0,000
Vald	-0,011506	0,017167	-0,670	0,503
Agri	-0,203308	0,045041	-4,510	0,000
Minas	0,230344	0,047752	4,820	0,000
industria	-0,069938	0,045228	-1,550	0,122
Elec	-0,004020	0,052682	-0,080	0,939
construc	-0,022143	0,045500	-0,490	0,627
Comer	0,009309	0,045041	0,210	0,836
Trans	0,049830	0,045694	1,090	0,275
Fin	0,128308	0,046278	2,770	0,006
Serv	-0,166334	0,044900	-3,700	0,000
_cons	10,624970	0,046083	230,560	0,000
select				
ocup	3,004522	0,032995	98,050	0,000
d_etnia	-0,412957	0,032088	-18,330	0,000
d_urbano	0,685433	0,024333	30,890	0,000
d_hombre	-0,047322	0,025189	7,170	0,001

numper	0,014445	0,007754	1,800	0,001
exper	0,015073	0,002245	6,520	0,000
exper2	-0,000431	0,000036	-15,370	0,000
arau	0,065877	0,075764	1,640	0,126
cau	-0,209733	0,032123	-19,630	0,000
jefe	0,405291	0,029912	21,780	0,000
n_hijo	-0,034021	0,010667	-2,790	0,000
_cons	-2,199972	0,054561	-38,870	0,000
<hr/>				
/athrho	0,188453	0,019066	9,880	0,000
/lnsigma	-0,402417	0,002715	-148,200	0,000
<hr/>				
Rho	0,186253	0,018405		
Sigma	0,668702	0,001816		
lambda	0,124548	0,012423		

Tabla A5 Ecuación del Salario 2009, MCO

Source	SS	df	MS	Number of obs =	85975
Model	14.009,6474	28	500,3446	F(28, 43538) =	1266,85
Residual	33.944,5406	85946	0,3950	Prob > F =	0,0000
Total	47.954,1880	85974	0,5578	R-squared =	0,2921
				Adj R-squared =	0,2919
				Root MSE =	0,6285

In_wage	Coef.	Std. Err.	t	P>t
exper	0,025708	0,000470	54,750	0,000
exper2	-0,000326	0,000008	-41,400	0,000
d_etnia	-0,053355	0,008060	-6,620	0,000
d_hombre	0,339651	0,005032	67,490	0,000
d_urbano	0,030051	0,005337	5,630	0,000
primc	0,102639	0,015018	6,830	0,000
seci	0,266637	0,016060	16,600	0,000
secc	0,480850	0,015805	30,420	0,000
teri	0,720604	0,018457	39,040	0,000
terc	1,246883	0,016849	74,000	0,000
ari	0,078371	0,024538	3,190	0,001
pari	-0,075010	0,047901	-1,570	0,117
arau	-0,109397	0,016527	-6,620	0,000
cau	-0,173866	0,010598	-16,410	0,000
malle	-0,181425	0,014916	-12,160	0,000
chil	0,016965	0,014347	1,180	0,237
osor	-0,094117	0,015555	-6,050	0,000
stgo	0,128035	0,006608	19,380	0,000
vald	-0,058734	0,015296	-3,840	0,000
agri	-0,158695	0,024827	-6,390	0,000
minas	0,355035	0,027826	12,760	0,000
industria	-0,030759	0,025292	-1,220	0,224
elec	0,012840	0,033589	0,380	0,702
construc	0,033091	0,025436	1,300	0,193
comer	-0,020091	0,024813	-0,810	0,418
trans	0,100710	0,025705	3,920	0,000
fin	0,090013	0,026298	3,420	0,001
serv	-0,097978	0,024698	-3,970	0,000
_cons	11,341730	0,029547	383,860	0,000

Tabla A6 Ecuación del salario 2009, Sesgo de Selección por Máxima Verosimilitud

Heckman selection model		Number of obs	=	101453	
(regression model with sample selection)		Censored obs	=	15478	
		Uncensored obs	=	85975	
		Wald chi2(28)	=	35475,0	
Log likelihood	=	-100831,40	Prob > chi2	=	0,0

In_wage	Coef.	Std. Err.	z	P>z
exper	0,026271	0,000474	55,440	0,000
exper2	-0,000337	0,000008	-42,260	0,000
d_etnia	-0,059573	0,008096	-7,360	0,000
d_hombre	0,341218	0,005039	67,710	0,000
d_urbano	0,034882	0,005368	6,500	0,000
primc	0,104136	0,015003	6,940	0,000
seci	0,268162	0,016046	16,710	0,000
secc	0,480949	0,015790	30,460	0,000
teri	0,721728	0,018443	39,130	0,000
terc	1,248403	0,016835	74,150	0,000
ari	0,075530	0,024522	3,080	0,002
pari	-0,077957	0,047833	-1,630	0,103
arau	-0,109811	0,016543	-6,640	0,000
cau	-0,178271	0,010618	-16,790	0,000
malle	-0,181559	0,014909	-12,180	0,000
chil	0,017229	0,014338	1,200	0,230
osor	-0,094233	0,015545	-6,060	0,000
stgo	0,127839	0,006607	19,350	0,000
vald	-0,058158	0,015291	-3,800	0,000
agri	-0,158588	0,024820	-6,390	0,000
minas	0,356057	0,027821	12,800	0,000
industria	-0,030450	0,025286	-1,200	0,229
elec	0,013669	0,033582	0,410	0,684
construc	0,033783	0,025430	1,330	0,184
comer	-0,020452	0,024806	-0,820	0,410
trans	0,101336	0,025699	3,940	0,000
fin	0,090246	0,026292	3,430	0,001
serv	-0,102445	0,024696	-4,150	0,000
_cons	11,322520	0,029610	382,390	0,000
select				
ocup	3,468500	0,023894	145,160	0,000
d_etnia	-0,298687	0,021369	-13,980	0,000
d_urbano	0,269021	0,014653	18,360	0,000
d_hombre	0,131241	0,015005	8,750	0,000

numper	0,016596	0,005743	2,890	0,004
exper	0,018073	0,001311	13,780	0,000
exper2	-0,000443	0,000021	-21,270	0,000
arau	-0,055851	0,050318	-1,110	0,267
cau	-0,226178	0,028384	-7,970	0,000
jefe	0,392667	0,017571	22,350	0,000
n_hijo	-0,023139	0,008436	-2,740	0,006
_cons	-2,259581	0,035509	-63,630	0,000
<hr/>				
/athrho	0,150004	0,016680	8,990	0,000
/lnsigma	-0,462875	0,002452	-188,790	0,000
<hr/>				
rho	0,148889	0,016310		
sigma	0,629472	0,001543		
lambda	0,093722	0,010311		

Tabla A7 Descomposición de Blinder – Oaxaca 1996

Blinder-Oaxaca decomposition				
1: d_etnia = 0				
2: d_etnia = 1				
In_wage	Coef.	Robust Std. Err.	z	P>z
Differential				
Prediction_1	11,652	0,004	2724,350	0,000
Prediction_2	11,271	0,017	655,880	0,000
Difference	0,381	0,018	21,500	0,000
Decomposition				
Explained	0,229	0,010	23,550	0,000
Unexplained	0,151	0,017	8,840	0,000

Tabla A8 Descomposicion de Blinder – Oaxaca 2000

Blinder-Oaxaca decomposition				
1: d_etnia = 0				
2: d_etnia = 1				
In_wage	Coef.	Robust Std. Err.	z	P>z
Differential				
Prediction_1	11,794	0,003	3.832,060	0,000
Prediction_2	11,444	0,012	994,260	0,000
Difference	0,350	0,012	29,390	0,000
Decomposition				
Explained	0,170	0,006	28,040	0,000
Unexplained	0,180	0,011	16,300	0,000

Tabla A9 Descomposición de Blinder – Oaxaca 2009

Blinder-Oaxaca decomposition				
1: d_etnia = 0				
2: d_etnia = 1				
In_wage	Coef.	Robust Std. Err.	z	P>z
Differential				
Prediction_1	12,332	0,003	4.651,120	0,000
Prediction_2	12,091	0,009	1.390,360	0,000
Difference	0,240	0,009	26,420	0,000
Decomposition				
Explained	0,140	0,004	31,790	0,000
Unexplained	0,101	0,008	11,930	0,000

Tabla A10 Estimación SAR 2009

Weights matrix	Number of obs	=	717
Name: pesos	Variance ratio	=	0,28
Type: Distance-based (binary)	Squared. Corr.	=	0,28
Distance band: 0,0 < d <= 2,5	Log likelihood	=	-666,08
Row-standardized: No	sigma	=	0,61

ln_wage	Coef.	Std. Err.	z	P>z
exper	0,025812	0,005285	4,88	0,000
exper2	-0,000310	0,000089	-3,49	0,000
d_etnia	-0,166344	0,093252	-1,78	0,074
d_hombre	0,287518	0,055490	5,18	0,000
d_urbano	0,152602	0,061696	2,47	0,013
primc	0,290552	0,172149	1,69	0,091
seci	0,431737	0,182032	2,37	0,018
secc	0,646939	0,178341	3,63	0,000
teri	0,914469	0,203203	4,50	0,000
terc	1,275847	0,193493	6,59	0,000
ari	0,295296	0,260764	1,13	0,257
pari	-0,680326	0,292943	-2,32	0,020
arau	-0,545156	0,182731	-2,98	0,003
cau	-0,175552	0,118165	-1,49	0,137
malle	-0,294129	0,175783	-1,67	0,094
chil	0,128279	0,153459	0,84	0,403
osor	-0,012724	0,160407	-0,08	0,937
stgo	0,043162	0,068005	0,63	0,526
vald	-0,324084	0,145487	-2,23	0,026
agri	-0,180677	0,316034	-0,57	0,568
minas	0,248927	0,347490	0,72	0,474
industria	-0,093452	0,318626	-0,29	0,769
elec	0,139201	0,387252	0,36	0,719
construc	-0,001538	0,319891	0,00	0,996
comer	-0,184456	0,313573	-0,59	0,556
trans	-0,009308	0,321388	-0,03	0,977
fin	-0,144615	0,332062	-0,44	0,663
serv	-0,289854	0,313435	-0,92	0,355
_cons	11,392000	0,366869	31,05	0,000

rho -0,000042 0,000017 -2,48 0,013

Wald test of rho=0 : chi2(1) = 6,145 (0,013)

Likelihood ratio test of rho=0 : chi2(1) = 6,119 (0,013)

Lagrange multiplier test of rho=0 : chi2(1) = 6,071 (0,014)

Acceptable range for rho: -1,000 < rho < 1,000

Tabla 12 Estimación por máxima verosimilitud 2009, con muestra reducida

ML model 2009			Number of obs	=	717
			Wald chi2(28)	=	270,58
Log likelihood	=	-669,1436	Prob > chi2	=	0,0
ln_wage	Coef.	Std. Err.	z		P>z
exper	0,025680	0,005308	4,840		0,000
exper2	-0,000309	0,000089	-3,470		0,001
d_etnia	-0,136039	0,092842	-1,470		0,143
d_hombre	0,285964	0,055724	5,130		0,000
d_urbano	0,154396	0,061955	2,490		0,013
primc	0,278701	0,172818	1,610		0,107
seci	0,432075	0,182810	2,360		0,018
secc	0,632494	0,179008	3,530		0,000
teri	0,905355	0,204039	4,440		0,000
terc	1,259137	0,194203	6,480		0,000
ari	0,416522	0,257232	1,620		0,105
pari	-0,581791	0,291475	-2,000		0,046
arau	-0,493519	0,182316	-2,710		0,007
cau	-0,106595	0,115336	-0,920		0,355
malle	-0,241978	0,175266	-1,380		0,167
chil	0,242198	0,147043	1,650		0,100
osor	0,053452	0,158846	0,340		0,736
stgo	0,038069	0,068265	0,560		0,577
vald	-0,240788	0,142158	-1,690		0,090
agri	-0,160874	0,317284	-0,510		0,612
minas	0,334678	0,347242	0,960		0,335
industria	-0,071260	0,319862	-0,220		0,824
elec	0,179460	0,388566	0,460		0,644
construc	0,022111	0,321116	0,070		0,945
comer	-0,142006	0,314444	-0,450		0,652
trans	0,001710	0,322732	0,010		0,996
fin	-0,109858	0,333185	-0,330		0,742
serv	-0,260161	0,314545	-0,830		0,408
_cons	11,245690	0,363639	30,930		0,000
rho					
_cons	-0,485684	0,026407	-18,390		0,000