UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN
FACULTAD DE ECONOMÍA
DIVISION DE ESTUDIOS DE POSGRADO

"LA IMPORTANCIA DE CARACTERISTICAS NO OBSERVABLES EN LA ESTIMACIÓN DE LOS DIFERENCIALES SALARIALES ENTRE INDUSTRIAS: ANÁLISIS PARA MÉXICO"

Por

Luz Irene Urbiola Domínguez

Tesis presentada como requisito parcial para obtener el grado de Maestría en Economía con Orientación en Economía Industrial

SEPTIEMBRE 2011

"LA IMPORTANCIA DE CARACTERISTICAS NO OBSERVABLES EN LA ESTIMACIÓN DE LOS DIFERENCIALES SALARIALES ENTRE INDUSTRIAS: ANÁLISIS PARA MÉXICO**"**

Luz Irene Urbiola Domínguez

Aprobación de Tesis:

Asesor de la Tesis **DR. ERICK RANGEL GONZÁLEZ**

DR. JORGE NOEL VALERO GIL

DR. JULIO CÉSAR ARTEAGA

DR. JULIO CÉSAR ARTEAGA
Director de la División de Estudios de Posgrado
De la Facultad de Economía, UANL
Septiembre, 2011.

DEDICATORIA

"Ama y haz lo que quieras. Si callas, callarás con amor; si gritas, gritarás con amor; Si corriges, corregirás con amor; si perdonas, perdonarás con amor. Si tienes el amor arraigado en ti, ninguna otra cosa sino el amor serán tus frutos".

San Agustín

A mis padres, Gonzalo y Anabel.

A Victor, el amor de mi vida.

AGRADECIMIENTOS

Agradezco a Dios que me ha permitido alcanzar esta meta.

A mis padres, que con su interminable amor han sido ejemplo de vida, y por sus consejos y paciencia, que me formaron dentro de los valores y principios sobre los se rige hoy mi vida.

A mi novio Victor, quien en todo momento ha estado a mi lado para ser el apoyo y la fuerza que me impulsan a conseguir mis metas, siempre con una palabra de aliento en los momentos difíciles y compartiendo mi alegría en los momentos felices.

A mis maestros, que con su esfuerzo me han dado el conocimiento que hoy rinde sus frutos al presentar este trabajo.

TABLA DE CONTENIDO

I. Introducción…….. 5

II. Revisión de la Literatura……..… 11

i. Estudios con resultados a favor de la teoría neoclásica de determinación salarial……...12

ii. Estudios con resultados a favor de la teoría de salarios de eficiencia……...13

iii. Discusión……..15

III. Modelo……..17

IV. Datos

i. Formación de la base de datos….…….23

ii. Variables……...25

iii. Análisis Descriptivo……..28

V. Estimaciones y Resultados

i. Categorización o '*Sorting*'……35

ii. Efectos Fijos……..……..…… 40

iii. Medida de Sesgo.….……..…….. 43

VI. Conclusiones……...…….46

VII. Bibliografía…….....48

VIII. Apéndice

i. Demostración 1…….50

ii. Definición de variables de la ENOE..……...51

I. INTRODUCCIÓN

De acuerdo a la teoría neoclásica, es la interacción de la oferta y la demanda entre empleadores y trabajadores lo que determina el salario de equilibrio, de modo que la existencia de desempleo sólo puede observarse si el trabajador lo decide voluntariamente. Entonces, una tasa natural de desempleo involuntario parece ser inconsistente con la teoría del mercado, puesto que existen personas dispuestas a trabajar al salario vigente pero no existen empleos para ellos.

Existen posiciones contradictorias con respecto al desempleo. Por un lado, la escuela neoclásica otorga un papel fundamental al mercado como el mejor mecanismo para asignar recursos, de modo que una tasa natural de desempleo puede reflejar ya sea que existe rotación de personal (desempleo friccional) y/o desajustes en la localización o cualificación requeridas para un empleo (desempleo estructural). Por otro lado, la teoría de salarios de eficiencia señala que el desempleo puede ser resultado de que algunos empleadores tienen incentivos a pagar un salario mayor al de mercado para compensar los efectos causados por la información imperfecta, ocasionando una oferta mayor a la demanda en el mercado de trabajo.

Cuando se observan diferenciales salariales entre individuos similares, estos pueden ser atribuibles a factores que no se pueden contabilizar en la asignación de la mano de obra o puede ser que se estén pagando salarios de eficiencia. Conocer el origen de estos diferenciales permite entender el mercado laboral y cómo se explica la tasa natural de desempleo. Es por esto que la brecha salarial entre industrias es algo que ha interesado a los economistas desde hace tiempo, en especial porque se ha encontrado evidencia de que hay un patrón similar para muchos países y que persiste en el tiempo[1].

Los modelos neoclásicos ofrecen una explicación a la existencia de diferenciales entre industrias incorporando la heterogeneidad de los individuos. Señalan que las industrias tienen características diferentes que se reflejan en la compensación al trabajador y también diferentes

[1]Véase Krueger (1988), Dickens y Katz (1987).

demandas de trabajadores, ya que algunas solicitan más años de educación, más experiencia laboral, etcétera, mientras que otras pueden no requerir ninguna calificación. Entonces, además de demandas por características observables, las industrias también tienen demandas por la habilidad de cada persona, es decir, por un capital humano que no es contabilizable pero que sí es compensado en el salario, por ejemplo: liderazgo, facilidad de palabra, capacidad de trabajo bajo presión, entre otros. La teoría de '*sorting*' dice que los individuos más hábiles se van a autoseleccionar y categorizar hacia las industrias de mayor pago (o mayor demanda) de modo que lo que parecería a simple vista ser un diferencial salarial, en realidad se podría explicar si pudiéramos controlar por las características de capital humano no observable de los individuos.

Los modelos de salarios de eficiencia señalan que las fallas en el mercado laboral asociadas a la información imperfecta causan que los empleadores paguen un salario mayor al de equilibrio para incentivar un comportamiento o resultado deseable por parte de los trabajadores, ya sea un mayor esfuerzo al trabajar o la permanencia en el empleo. Katz (1986) hace una revisión de estos modelos y los clasifica en cuatro tipos:

1. El primer modelo es el de '*shirking'* (elusión), propuesto por Shapiro y Stiglitz (1982). El supuesto principal es que no se puede observar perfectamente el nivel de esfuerzo del trabajador ni su desempeño, y monitorearlo es costoso; de aquí que los beneficios de aplicar un salario mayor provienen de generar un costo al trabajador de perder su empleo, y lo incentivan a esforzarse más, lo que permite ahorrar a la empresa en costos de monitoreo.
2. El segundo modelo es el de 'turnover'[2] (rotación), el cual asume que las empresas cargan con parte de los costos de contratación y/o entrenamiento. Por lo tanto, las empresas desean tener la menor rotación posible, de modo que aumentan el costo de oportunidad de renunciar para el trabajador pagándole un mayor salario que el de mercado. De este modo, el trabajador tomará en cuenta que perder el trabajo también implica perder un mayor salario (incluso superior al valor de su productividad).

[2] Este modelo es muy similar en su estructura al de 'shirking'. Véase por ejemplo Salop (1979) o Stiglitz (1974).

3. Un tercer modelo es el de selección adversa[3] que asume que existen diferentes tipos de trabajadores y que no se puede observar perfectamente la calidad y el futuro desempeño del trabajador al momento de contratarlo, por lo que la empresa busca ampliar la posibilidad de contratar a trabajadores de alta calidad empleando una estrategia de ofrecer salarios más altos para atraerlos. Un salario de eficiencia significaría que se paga un salario más alto para tener la opción de elegir entre varios solicitantes para obtener al mejor posible.
4. El cuarto modelo es el de sindicatos, en el que el supuesto principal es que los costos de reemplazo de la fuerza laboral actual otorgan poder de negociación a los trabajadores, de tal forma que con mayores salarios se evita la sindicalización y se mantiene la paz industrial. Y por último, también desde un punto de vista sociológico, salarios altos incentivan mayor lealtad hacia la empresa y mejor moral.

Existe una amplia literatura que analiza de manera empírica la posible existencia de diferenciales salariales entre industrias[4] para Estados Unidos y otros países desarrollados, no obstante, hay pocos estudios que se enfoquen en México. Sin embargo, la cercanía a Estados Unidos así como los grandes cambios que ha sufrido en sus políticas comerciales hacen del mercado laboral mexicano un interesante objeto de estudio.

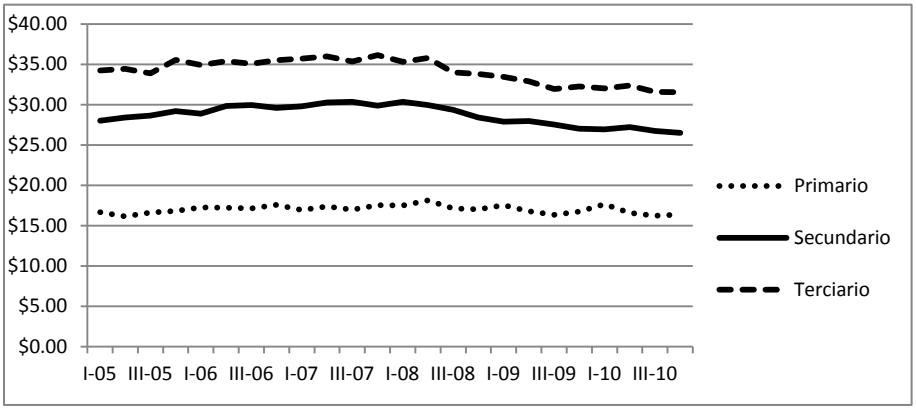
La estructura salarial en México ha mostrado marcados diferenciales a lo largo del tiempo, los cuales se acentuaron principalmente por la implementación de políticas de liberalización comercial que afectaron la organización y funcionamiento de los mercados y, por lo tanto, las funciones de oferta y demanda del mercado laboral[5].

En los últimos años, los diferenciales salariales se han mantenido relativamente estables entre las industrias, e incluso durante las épocas de recesión económica no se ha observado una tendencia significativa de ampliación o reducción de la brecha salarial. En el gráfico 1 se puede observar que entre 2005 y 2010 los salarios por hora en los sectores secundario y terciario tienen el mismo

[3] Véase por ejemplo Weiss (1980) o Stiglitz (1976)
[4] La literatura sobre los diferenciales salariales se estudia en el siguiente capítulo.
[5] Véase Alarcón y McKinley (1997).

comportamiento, manteniéndose una diferencia entre los salarios de los dos sectores aparentemente igual.

Gráfico 1: Salario por hora real. Por sector de la economía. México



Fuente: Elaboración Propia con datos de las ENOE 2005-2010. Pesos constantes a precios de diciembre de 2010.

Aunque el salario por hora en el sector primario no sigue la misma tendencia que los otros dos, la brecha entre los sectores primario y terciario y entre los sectores primario y secundario se mantuvo sin variaciones importantes entre 2005 y 2008, y luego tuvieron una leve reducción a partir del tercer trimestre de 2008, época en la que el país estaba en recesión económica. De acuerdo a la teoría de salarios de eficiencia, las empresas en industrias competitivas tienden a ser más elásticas en los pagos ante una crisis económica o recesión, mientras que las empresas en industrias que pagan salarios por encima al de equilibrio pueden reducir el salario pero en una menor proporción[6]. Sin embargo, no se puede aseverar a partir del gráfico que se paguen salarios de eficiencia en algún sector. Los diferenciales pueden ser ocasionados por diversos factores, desde diferentes demandas de trabajo calificado en cada empresa hasta compensaciones por arriba del salario de equilibrio para reducir la rotación de personal.

A diferencia de los estudios para México que se han realizado hasta ahora que se han enfocado en análisis teórico y cualitativo de los mercados laborales, esta investigación pretende aportar evidencia empírica sólida acerca de la estructura salarial en México y determinar si los diferenciales entre industrias que se observan y persisten son el resultado de compensaciones salariales a

[6] Véase Katz, L. (1986)

características individuales no observadas o si se debe a que se pagan salarios de eficiencia en algunas industrias. La hipótesis principal es que los diferenciales salariales entre industrias reflejan en buena parte que existe categorización de los individuos a través de sus características no observables como el nivel de habilidad.

En el siguiente capítulo se hace una revisión de la literatura empírica sobre los diferenciales salariales entre industrias y la principal aportación de cada estudio al campo de la economía laboral. Destacan artículos como el de Gibbons y Katz (1992), en el que se desarrolla un modelo de movilidad interindustrial endógena para mostrar porqué es posible que existan diferenciales salariales en equilibrio. Explican cómo las diferentes demandas y un proceso de *"matching"* pueden llevar a que los individuos se coloquen en la industria que demanda su nivel de habilidad y que la compensa por medio de un salario más alto en presencia de información imperfecta. También destaca el artículo de Murphy y Topel (1990) ya que utilizan un modelo diferente para contrastar la hipótesis de categorización ('*sorting*') como posible explicación de los diferenciales salariales. Prueban que existen características individuales y de capital humano observable que acomodan a los individuos en las industrias que mejor les compensan sus propias características. Utilizan una metodología que permite identificar el efecto de categorización y hacer una aproximación a la proporción de sesgo que existe a causa de la omisión de variables no observables bajo ciertos supuestos.

En el tercer capítulo se describe el modelo teórico propuesto por Combes, Duranton y Gobillon (2005) en el cual las diferencias salariales que se observan son causadas por características de la industria, zona laboral y la habilidad del trabajador. Después se describe la metodología econométrica que se utiliza para contrastar la hipótesis de categorización a través de características observables, que asume que ese mismo proceso que se da a través del capital humano no observable. A continuación se describe la metodología de efectos fijos para eliminar las características no observadas individuales y después, el cálculo que se utiliza para estimar el sesgo causado en las estimaciones debido a la omisión de la habilidad en la ecuación de salario.

El capítulo cuarto describe la manera en que se manipularon las encuestas para construir la base de datos, la cual consta de las Encuestas Nacionales de Ocupación y Empleo (ENOE) desde 2005 hasta 2010 y es un panel no balanceado de individuos. En este capítulo también se describen las variables importantes a utilizar y se hace un análisis descriptivo de los datos.

En el quinto capítulo se discuten los resultados así como su implicación para el objetivo de este trabajo. Se encontró evidencia de existe categorización ('*sorting')* entre los trabajadores de acuerdo a los años de educación, variable que es indicador de mayor capital humano, y se encontró evidencia de que el capital humano no observable puede explicar una gran parte de los diferenciales salariales entre industrias. Por último en la sexta parte se listan las conclusiones.

II. REVISIÓN DE LA LITERATURA

Existe una extensa variedad de artículos dedicada a estudiar los determinantes de los diferenciales salariales entre industrias. La discusión de la teoría económica y los datos disponibles han dado lugar a diversas posturas: mientras que algunos autores han encontrado evidencia de verdaderos diferenciales salariales[7], es decir, que se pagan salarios de eficiencia en algunas industrias, hay otros que consideran que esas conclusiones no son correctas puesto que si se pudiera controlar por otros efectos no observados, los diferenciales desaparecerían[8].

Artículos como el de Gibbons y Katz (1992) o el de Combes, Duranton y Gobillon (2005) desarrollan un modelo de determinación salarial neoclásico, en el que se asume información perfecta por lo que el salario del trabajador i es igual al valor de su producto marginal. En estos modelos la variación salarial entre trabajadores proviene ya sea de sus características individuales (educación, habilidad, experiencia) o de las características de la industria o zona geográfica donde se localice la empresa.

Por otro lado, Daron Acemoglu (1998) propone un modelo teórico en el que los diferenciales salariales dependen del nivel de capital humano de los individuos y varían a medida que se da el cambio tecnológico, puesto que los avances tecnológicos suelen estar diseñados para utilizarse por las personas con mayor capital humano o mayor habilidad. En este sentido, los diferenciales entre industrias estarían explicados por diferencias en las demandas por capital humano y el uso de tecnologías.

Lawrence Katz (1986) hace una revisión de la literatura en donde explica los principales modelos de salarios de eficiencia y las aplicaciones de esta teoría, entre las que se encuentra la de explicar los diferenciales salariales entre industrias. Los modelos están basados en el supuesto de que existe información imperfecta en el mercado laboral por lo que los incentivos del mercado

[7] Autores como Dickens y Katz (1987), Krueger y Summers(1986), Krueger y Summers(1988) encuentran evidencia consistente de salarios de eficiencia basados.
[8] Murphy y Topel (1989), Blackburn y Neumark (1991)

llevan a un salario mayor al de equilibrio. Shapiro y Stiglitz (1982) propusieron el modelo de elusión o *'shirking'* en el cual existe un equilibrio con desempleo puesto que si todas las empresas siguen la misma estrategia, el salario de equilibrio sería mayor al salario de mercado y habría una cantidad de personas que quedarían sin trabajo y que incluso estarían dispuestos a trabajar por menos dinero, pero debido que está la posibilidad de que holgazaneen, no pueden prometer no hacerlo de manera que sea creíble.

i. Estudios con resultados a favor la teoría neoclásica de determinación salarial.

Murphy y Topel (1990) analizan los diferenciales salariales para individuos que cambiaron de industria u ocupación y prueban la hipótesis de que se da un proceso de autoselección a través características observables del individuo. Dicha hipótesis incorpora la heterogeneidad de los individuos y señala que las industrias tienen diferentes demandas de trabajadores, lo que significa que sus demandas varían tanto en las características observables como no observables. Por lo tanto, los individuos se acomodan entre industrias y ocupaciones a través de sus características como la edad, años de educación, estado civil, etc. (observables) y habilidad (capital humano no observables).

El artículo desarrolla una metodología para identificar el efecto de categorización de los individuos utilizando un panel construido a partir de la encuesta Current Population Survey March (CPS March) desde 1977 a 1983 controlando por variables de la industria, de la ocupación y de características del trabajador y estiman dos regresiones auxiliares para probar la hipótesis de '*sorting'*. Sus resultados sugieren que existe categorización a través de observables y que los individuos que cambian industria u ocupación sólo reciben una tercera parte del diferencial salarial estimado por corte transversal y concluyen que sólo una pequeña parte de los diferenciales salariales no puede ser explicada por el sesgo de omitir la habilidad.

Blackburn y Neumark (1992) estiman los diferenciales salariales por medio de variables instrumentales utilizando el puntaje en pruebas de inteligencia como proxy de la habilidad del

individuo. Obtienen la mayor parte de sus datos de la encuesta National Longitudinal Survey (NLS) de 1966 para individuos que vuelven a ser entrevistados en 1973 y 1980. Sus resultados indican que aunque la habilidad sí explica parte de los diferenciales, sólo representa una pequeña parte.

Combes, Duranton & Gobillon (2005) hacen un estudio utilizando un panel de trabajadores franceses los cuales fueron monitoreados desde 1976 hasta 1998. Controlan por efectos fijos del trabajador, área, año e industria y sus resultados sugieren que la habilidad individual contabiliza una gran parte de las disparidades espaciales, especialmente evidencia de categorización de los individuos entre regiones a través de su habilidad. Keane (1991) estudia el mercado laboral estadounidense utilizando un panel de trabajadores de dieciséis años y utiliza efectos fijos para controlar por la habilidad no observada de los individuos. Encuentra que no hay evidencia que apoye la teoría de salarios de eficiencia. Además en este estudio, también analiza los ciclos económicos para ver si existe alguna divergencia entre las industrias que se piensa pagan salarios de eficiencia y las que no. Encuentra que la brecha salarial es estable incluso durante las crisis, por lo que afirma en sus resultados que no existen los salarios de eficiencia.

ii. Estudios con resultados a favor de la teoría de salarios de eficiencia.

Dickens y Katz (1987) hacen un estudio utilizando datos de la CPS empleando todos los trimestres de 1983 en el que encuentran que aún después de controlar por la características de la industria, éstas variables solamente explican el 6.7% de los diferenciales salariales interpersonales y a lo mucho un 30% en total. También encuentran que el tamaño de la empresa y la relación capital-trabajo tienen una relación positiva significativa en la determinación del salario.

Krueger y Summers (1986) encuentran que los patrones de diferencias salariales son similares en el tiempo, entre países, y entre ocupaciones. También analizan los diferenciales por medio de variables como el poder de mercado de la empresa y la razón capital-trabajo y concluyen que los modelos competitivos no permiten explicar el mercado laboral y que la evidencia se ajusta más a las teorías de salarios de eficiencia. Dos años más tarde, Krueger y Summers (1988) realizan un estudio

con datos de corte transversal y de panel para probar la hipótesis de salarios de eficiencia. Encuentran que existe una gran dispersión de los salarios entre las industrias, medida como la desviación estándar en diferenciales salariales en la industria. Ellos usan el método de efectos fijos para eliminar el sesgo generado por la habilidad y otras características no observadas de los trabajadores y corrigen los errores para aquellos trabajadores que cambian de industria. Encuentran que aún controlando por efectos fijos individuales, características del trabajador, ocupación e industria, persisten los diferenciales en los salarios y atribuyen este resultado a que las empresas comparten sus rendimientos a los trabajadores para incentivarlos a cambiar su conducta en beneficio de la empresa.

Gibbons y Katz (1992) hacen un estudio en el cual enfrentan la hipótesis de que en realidad existen diferenciales salariales entre las industrias contra la hipótesis de que los diferenciales reflejan las habilidades productivas de los trabajadores que no se pueden medir. Utilizan el CPS de 1984 y 1986 para trabajadores que cambiaron de trabajo debido a cierres de plantas. Encuentran que no se explican por completo los diferenciales por las variables de control debido a que las diferencias salariales están correlacionadas con las ocupaciones, es decir, en las industrias donde se paga mucho para alguna ocupación, también tiende a serlo para las demás ocupaciones. No encuentran evidencia de categorización ('*sorting*') y al utilizar primeras diferencias encuentran que los diferenciales son del mismo signo y con magnitud similar. Coinciden con Krueger y Summers en que existen rentas en las industrias de mayores diferenciales salariales positivos que reducen la rotación del personal.

Alarcón y McKinley (1997) realizan un estudio sobre la desigualdad salarial y principalmente sobre el cambio en la estructura salarial durante los años 1989-1992 y observan que la brecha salarial entre trabajadores calificados y no calificados se amplió, y que los salarios para diferentes ocupaciones también varió, esto como consecuencia de la implementación de políticas comerciales importantes que afectaron la oferta y demanda laboral.

Salas & García (2010) presentan un resumen de las posibles explicaciones por teorías de salarios de eficiencia del desempleo en México. Salas & Zepeda (2003) hacen un estudio descriptivo de la estructura salarial entre industrias para México de 1990 a 2001, en el cual encuentran evidencia de que las reformas económicas y los programas de estabilización tuvieron impacto sobre los diferenciales salariales, ampliando la brecha entre los trabajadores calificados y no calificados, y entre la industria manufacturera (maquiladoras) con respecto a las demás.

iii. Discusión

Aunque hay varios trabajos apoyando ambas teorías, la evidencia no es concluyente. Todavía hay cuestiones acerca de la metodología que dan lugar a diversas opiniones, en especial las referentes al efecto de la habilidad en la determinación del salario. Para corregir este sesgo en la estimación de los diferenciales salariales existen métodos econométricos como el uso de variables instrumentales, que en teoría incorpora una estimación de la variable omitida por medio de otras variables llamadas "instrumentos" que no estén correlacionados con el error. Sin embargo, lo complicado es encontrar los instrumentos adecuados que calculen de manera correcta la habilidad de los individuos, ya que de otro modo el sesgo no se reduce.

Otra metodología que se usa con frecuencia es la de estimar los diferenciales salariales por medio de efectos fijos o primeras diferencias. Sin embargo, los datos de panel suelen ser difíciles de obtener para un período lo suficientemente largo como para que los diferenciales no representen un efecto del ciclo económico. También es posible que la habilidad no observada en los individuos no se comporte como un efecto fijo en el tiempo, y el uso de efectos fijos o primeras diferencias no eliminaría el sesgo por habilidad de los diferenciales salariales.

La metodología a utilizar en este estudio es similar a la utilizada por Murphy y Topel (1990) para contrastar la hipótesis de categorización (*'sorting'*) por características observables. Asimismo, se emplea una segunda metodología, similar a la de Gibbons y Katz (1992) y Keane (1991) que estiman los diferenciales salariales entre industrias utilizando el método de efectos fijos.

Adicionalmente, la presente investigación también calculará una medida estimada del sesgo causado por la omisión de características no observables que afectan la productividad, como en Murphy & Topel (1990). Este sesgo se calcula bajo el supuesto de que los diferenciales industriales son cero y que el efecto del diferencial proviene de un efecto fijo que afecta el salario y da lugar a la autoselección hacia industrias de mayor pago, de modo que el mayor pago en las industrias vendría determinado únicamente por una mayor demanda de habilidad, y los individuos más hábiles se acomodarían en tales industrias, recibiendo cada trabajador en la economía un pago igual al valor de su productividad marginal.

III. MODELO

La teoría de categorización asume que la productividad de los individuos no se puede observar perfectamente y que está correlacionada con el nivel de educación[9]. Esta teoría mantiene los supuestos neoclásicos de maximización de los beneficios. La presente investigación sigue el modelo de Combes, Duranton y Gobillon (2005) de determinación salarial. Se asumen empresas competitivas idénticas dentro de cada industria y zona urbana[10], cada una conformada por individuos con un nivel de habilidad heterogéneo.

Considere una empresa representativa competitiva que opera en la industria k, en la zona urbana j en el momento t. Su función objetivo es:

$$\pi_{k,j,t} = p_{k,j,t} y_{k,j,t} - \sum_{i \in k,j,t} w_{i,t} l_{i,t} - r_{k,j,t} z_{k,j,t} \qquad (1)$$

Donde $p_{k,j,t}$ y $y_{k,j,t}$son el precio y la producción respectivamente. El salario del trabajador i y las horas trabajadas se denotan por $w_{i,t}$ y $l_{i,t}$ respectivamente; $r_{k,t} z_{k,t}$ es el valor de los demás factores de producción usados en la industria k en el trimestre t. Sea la función de producción de la empresa:

$$y_{k,j,t} = A_{k,j,t} \left(\sum_{(i \in k,j,t)} s_{i,t} l_{i,t} \right)^{\beta} \left(z_{k,j,t} \right)^{1-\beta} \qquad (2)$$

Donde el 0<β≤1, s_{it} es la habilidad del trabajador i en el momento t y $A_{k,j,t}$ es la productividad total de factores. En el equilibrio competitivo, el salario que recibe el trabajador i en el año t, que trabaja en la industria k debe ser igual a su productividad marginal. Y si sustituimos las condiciones de primer orden para la maximización de beneficios obtenida de la ecuación 1, obtenemos:

$$w_{it} = \beta (1-\beta)^{\frac{1-\beta}{\beta}} \left(p_{k(i,t),j(i,t),t} \frac{A_{k(i,t),j(i,t),t}}{r_{k(i,t),j(i,t),t}{}^{1-\beta}} \right)^{\frac{1}{\beta}} \cdot s_{i,t} \qquad (3)$$

$$w_{it} = B_{k(i,t),j(i,t),t} \cdot s_{i,t}$$

[9] Principalmente se supone que existe correlación entre la productividad y el nivel de educación, pero también se puede suponer que se correlaciona con otras variables observables de capital humano, como la experiencia.

[10] El término zona urbana se refiere a la cantidad de población, es decir, una zona se considera urbana si tiene 100,000 habitantes o más y rural si tiene menos.

La ecuación (3) representa el salario del trabajador i en el momento t. Note que la productividad de la industria, el precio del producto e incluso el precio de los demás factores de producción afectan el salario, de modo que es importante controlar por estos efectos.

Para poder estimar econométricamente la ecuación (3) con los datos disponibles, la transformaremos en una ecuación minceriana[11]. El factor $B_{k,j,t}$ es el que determina diferencias en los salarios para un nivel de habilidad dado. Asuma que dicho factor está dado por:

$$\log B_{k,j,t} = \delta_{k,t} + \gamma_{j,t} \tag{4}$$

Donde $\delta_{k,t}$ es un efecto fijo de la industria-trimestre y $\gamma_{j,t}$ es un efecto fijo de la zona-trimestre poblacional. Sin embargo, para la estimación econométrica se hará el supuesto de que los efectos de industria y zona son fijos, y se controlará por el período de tiempo de modo que la ecuación (4) se transforma en:

$$\log B_{k,j,t} = \delta_k + \gamma_j + t \tag{4a}$$

Donde t es un vector de coeficientes de efectos trimestre.

Asumiendo que la habilidad del individuo se determina tanto por factores observados y por factores no observados que son fijos en el tiempo:

$$\log s_{i,t} = X_{i,t}\beta + \alpha_i + \epsilon_{i,t} \tag{5}$$

Combinando la ecuación (3), (4a) y (5) tenemos:

$$\log w_{i,t} = \delta_k + \gamma_j + t + X_{i,t}\beta + \alpha_i + \epsilon_{i,t} \tag{6}$$

La ecuación (6) representa la ecuación de los determinantes del salario controlando por variables de la industria, zona urbana/rural, trimestre-año y características del trabajador. Entonces, la forma econométrica de la ecuación (6) es:

$$\log w_{i,t} = \delta_k Ind_{it} + \gamma_j Zona_{it} + \rho_t Trim + X_{i,t}\beta + \alpha_i + \epsilon_{i,t} \tag{6a}$$

Los coeficientes δ_k, γ_j, ρ_t y algunos elementos del vector β pueden ser estimados incluyendo variables dummy para la industria, la zona y algunas características observables del individuo. No

[11] La ecuación minceriana modela el salario en función del capital humano. Véase Mincer (1974).

obstante, no se cuenta con datos sobre el nivel de habilidad de los trabajadores para estimar el coeficiente α_i por lo que la estimación por mínimos cuadrados ordinarios de la ecuación (6a) dará como resultado coeficientes sesgados y por lo tanto no confiables de los diferenciales salariales.

Para contrastar la hipótesis de categorización, se seguirá la metodología utilizada por Murphy y Topel (1990). La idea central de esta hipótesis es que las industrias tienen diferentes demandas por algunas características del individuo, como aquellas que distinguen a los trabajadores por su productividad. Por ejemplo, algunas industrias tienen demandas por trabajadores con mayores años de educación debido a que el trabajo que realizarán requiere mayor capacitación o mayor nivel de conocimientos, mientras que en otras puede que esto no sea tan importante. Por otro lado, hay industrias que requieren ciertas habilidades como capacidad gerencial mientras que en otras no es valorado en el salario. La hipótesis de categorización busca probar que en realidad existen diferentes demandas entre las industrias por características de capital humano, principalmente por escolaridad. Para ello, es necesario observar los determinantes del salario y los efectos industria. Se estimará la ecuación (6a) por mínimos cuadrados ordinarios:

$$\log w_{i,t} = \mu + \delta_k Ind_{it} + \gamma_j Zona_{it} + \rho_t Trim + X_{i,t}\beta + e_{i,t} \qquad (7)$$

El parámetro μ es la constante, y el error $e_{i,t} = \alpha_i + \epsilon_{i,t}$. Para probar la hipótesis de categorización (sorting), se estiman dos regresiones auxiliares:

$$\log w_{i,t} = \mu_1 + X_{i,t}\beta_1 + u_{i,t} \qquad (8)$$

$$\text{efecto_total}_{\text{i,t}} = \mu_2 + X_{i,t}\beta_2 + v_{i,t} \qquad (9)$$

La ecuación (8) muestra en qué medida las características observables de los trabajadores determinan el salario, independientemente de la industria u ocupación en que trabajen. La ecuación (9) explica la relación que existe entre las características observables y el efecto salarial de la industria. Si la hipótesis de '*sorting*' es correcta, entonces existe una correlación entre X y la industria y zona, lo que significa que se da un proceso de autoselección de los trabajadores con mejores características observables hacia las industrias y zonas con mayores salarios. Para

confirmar la hipótesis, es necesario encontrar que los coeficientes de las variables explicativas de las ecuaciones (8) y (9) tengan el mismo signo.

La variable efecto total es calculada a partir de los coeficientes estimados en la ecuación (7). Es decir, el efecto total comprende el efecto salarial de la industria k en el momento t y de la zona rural/urbana en el momento t.

Puesto que los coeficientes estimados por MCO en la ecuación (7) están sesgados debido a que el error de la estimación contiene información que no puede ser incluida, se utilizará una metodología de efectos fijos similar a la utilizada por Keane (1991). Se busca controlar por las características individuales no observadas que afectan la productividad de los individuos y por lo tanto afectan el salario real. La metodología de efectos fijos se calcula por medio de una transformación en la cual se estima la media de la ecuación (6a) a través del tiempo para cada individuo:

$$\overline{\log w_i} = \bar{\mu} + \delta_k \overline{Ind}_i + \gamma_j \overline{Zona}_i + \bar{X}_i \beta + \bar{\alpha}_i + \bar{\epsilon}_i \qquad (10)$$

Y posteriormente, se calcula la media total de cada variable, es decir, la media a través del tiempo y a través de los individuos:

$$\overline{\overline{\log w}} = \bar{\bar{\mu}} + \delta_k \overline{\overline{Ind}} + \gamma_j \overline{\overline{Zona}} + \bar{\bar{X}} \beta + \bar{\bar{\alpha}} + \bar{\bar{\epsilon}} \qquad (11)$$

A partir de las ecuaciones (7), (10) y (11), se obtiene:

$$\log \ddot{w}_{it} = \mu + \delta_k \ddot{Ind}_{it} + \gamma_j \ddot{Zona}_{it} + \ddot{X}_{it} \beta + \bar{\bar{\alpha}} + \ddot{\epsilon}_{it} \qquad (12)$$

Donde: $\log \ddot{w}_{it} = \log w_{it} - \overline{\log w_i} + \overline{\overline{\log w}}$

$\mu = \mu - \bar{\mu} + \bar{\bar{\mu}}$ $\quad \bar{\bar{\alpha}} = \alpha - \bar{\alpha} + \bar{\bar{\alpha}}$ $\quad \ddot{\epsilon}_{it} = \epsilon_{it} - \bar{\epsilon}_i + \bar{\bar{\epsilon}}$

$\ddot{Ind}_{it} = Ind_{it} - \overline{Ind}_i + \overline{\overline{Ind}}$ $\quad \ddot{Zona}_{it} = Zona_{it} - \overline{Zona}_i + \overline{\overline{Zona}}$ $\quad \ddot{X}_{it} = X_{it} - \bar{X}_i + \bar{\bar{X}}$

Se utiliza el software estadístico Stata, que realiza la estimación de la ecuación (12) cuando se utiliza un modelo de efectos fijos bajo la restricción $\bar{\bar{\alpha}} = 0$, que se cumple si α_i varía entre industrias pero no a través del tiempo. Entonces, en este modelo el coeficiente μ representa el efecto fijo medio de todos los individuos. Bajo esta restricción, el error en la ecuación (12) no está

correlacionado con las variables independientes, de modo que los estimadores de MCO serán consistentes y eficientes.

Si la hipótesis principal es correcta y los diferenciales salariales se explican porque existen características no observadas de los individuos que se correlacionan con las industrias, entonces al utilizar el método de efectos fijos los coeficientes van ser significativamente menores a los estimados por la ecuación (7) o no significativos. Esto sería un indicador de que los efectos de la industria son iguales entre sí una vez que se controla por características del individuo observables y no observables y por la zona urbana.

Para complementar los resultados anteriores se calculará una medida de sesgo como la propuesta por Murphy y Topel (1990). Se estima el porcentaje del sesgo en la estimación de corte transversal que puede ser atribuido a la omisión de la habilidad y otras características no observadas del individuo. Partiendo del supuesto de que los efectos salariales de la industria son iguales a cero e incorporando el efecto fijo individual en la ecuación (7) se obtiene:

$$\log w_{i,t} = \mu_1 + X_{i,t}\beta_1 + \alpha_i + \varepsilon_{i,t} \tag{13}$$

El efecto total salarial de la industria sería, por otro lado, igual a:

$$\text{efecto_total}_{\text{i,t}} = \mu_2 + X_{i,t}\theta + \varphi\alpha_i + \varepsilon_{i,t} \tag{14}$$

La ecuación (13) representa la estimación de corte transversal de los determinantes del salario cuando los efectos de industria, zona y trimestre son cero. Es decir, supone que individuos observablemente similares reciben la misma compensación económica independientemente de la zona e industria en la que laboren y que no hay diferencias significativas reales en la remuneración de un trimestre a otro. La ecuación (14) representa la relación que existe entre los efectos salariales de la industria y las características del individuo y la habilidad. Aquí φ representa el efecto de la habilidad sobre el efecto salarial, o bien, la magnitud del efecto de autoselección para el individuo. Es decir, mientras mayor sea el coeficiente φ o mayor sea α, más fuerte será el efecto de categorización o ‘sorting’ entre industrias.

Cuando se hace la estimación de (7), el sesgo que existe en los coeficientes del efecto salarial de la industria es igual a la correlación entre el efecto industrial real y la habilidad multiplicado por la varianza de la variable omitida (α) entre la varianza del error de la regresión[12]. Por lo tanto, el sesgo se calcula de la siguiente manera:

$$SESGO = \varphi\sigma_{\alpha}^{2}/\sigma_{\varepsilon}^{2} \qquad (15)$$

Si la evidencia muestra que el sesgo representa una buena parte de los diferenciales salariales entre industrias, entonces se puede concluir que existe evidencia a favor de la teoría neoclásica de determinación salarial y que los coeficientes observados en la estimación de corte transversal pueden ser atribuidos al efecto de categorización entre los individuos de acuerdo a su habilidad.

[12] En el apéndice se incluye la demostración de que la ecuación (15) representa el sesgo.

IV. DATOS

i. Formación de la Base de Datos

La presente investigación emplea datos de la Encuesta Nacional de Ocupación y Empleo. La ENOE es una encuesta que consta de cinco bases de datos por trimestre con información sobre el hogar e información sobre los individuos que residen en el hogar y es una base de panel en donde se realizan entrevistas a los miembros del hogar durante 5 trimestres, pero en donde se va rotando al 20% de la muestra cada trimestre, de modo que los cortes transversales van formando un panel no balanceado como se muestra en la figura 1. Para este trabajo, se utilizan la base de datos de características sociodemográficas de los individuos del hogar, además de dos bases que incluyen el cuestionario sobre ocupación y empleo.

La base de características sociodemográficas incluye a todos los individuos del hogar, pero las bases de los cuestionarios únicamente incluyen a individuos mayores de 14 años[13].

Figura 1. Panel no balanceado de la ENOE

	I-05	II-05	III-	IV-	I-06	II-06	III-	IV-	I-07	II-07	III-	IV-	I-08	II-08	III-	IV-	I-09	II-09	III-	IV-	I-10	II-10	III-	IV-
Número de entrevista	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4
	5																				1	2	3	
	4	5																					1	2
	3	4	5																					1
	2	3	4	5																				5

Fuente: Elaboración Propia.

Se cuenta con un total de 24 encuestas desde el primer trimestre de 2005 hasta el cuarto trimestre de 2010. En cada levantamiento se recopila información de alrededor de 300 mil individuos, aunque en cada trimestre se pierde o no se puede volver a contactar a un 6% aproximadamente (INEGI, Estadísticas de la dinámica laboral en México 2005-2007, 2009).

Para poder unir las bases e identificar a los individuos, se generó un folio (identificador único) del individuo en cada período similar al que se sugiere en los manuales de ENOE, pero además se le

[13] En algunas encuestas se toma en cuenta a personas mayores de 12 años, de modo que para mantener el mismo criterio al unir la información, se tomó como población económicamente activa a las personas de 14 años o más.

agrega un control por fecha de nacimiento para evitar duplicados debido a la gran cantidad de observaciones (INEGI, Conociendo la base de datos de la ENOE, 2007). Este identificador consta de diez campos de información, siete de los cuales conforman la llave para unir la base de características sociodemográficas y los cuestionarios, y las tres últimas se incorporan a la llave para unir los trimestres entre sí[14].

$$Unico = cd_a + ent + con + n_{pro_{viv}} + v_{sel} + n_{hog} + h_{mud} + n_{ren}$$

$$Folio = Unico + Año_{nacimiento} + Mes_{nacimiento} + Día_{nacimiento}$$

Para la construcción de la base de datos a utilizar se consideraron hombres de 14 a 65 años que se encuentran ocupados y que reciben y reportan un pago por su trabajo al momento de la encuesta. La población no ocupada o los trabajadores que no reportan un salario se excluyen de la muestra. De estos individuos se eliminaron a aquellos de quienes no se tiene información suficiente o no la reportan correctamente como para clasificarlos dentro de una industria o que no se puede calcular alguna de sus características observables, como años de educación o experiencia. En la tabla 1 se observa que, una vez eliminadas las observaciones con respuestas no válidas se cuenta con un total de 1,277,155 observaciones correspondientes a 562,398 hombres de 14 a 65 años.

Tabla 1. Número de observaciones y entrevistas de la base de datos.

	Hombres		**Mujeres**	
Entrevistas	No. Obs.	Individuos	No. Obs.	Individuos
1	217004	217004	165424	165424
2	270728	135364	184408	92204
3	286698	95566	194559	64853
4	278380	69595	199068	49767
5	224345	44869	172335	34467
Total	1277155	562398	915794	406715

Fuente: Elaboración Propia.

Las principales ventajas de esta base de datos son que cuenta con una gran cantidad de observaciones y que abarca un periodo de tiempo relativamente amplio, lo que permite una mayor robustez en los estimadores y además permite descartar que los resultados reflejen factores

[14] El significado de los campos para la generación de la llave se incluye en el apéndice.

exógenos del ciclo económico[15]. Entre las desventajas se encuentra que la encuesta se aplica a hogares y no a establecimientos, de modo que podría existir un grado de inexactitud en las respuestas laborales de los trabajadores. También, el 6% del total de los entrevistados no se vuelven a contactar puesto que los encuestadores regresan a la misma vivienda y no se sigue a los miembros del hogar.

ii. Variables

La unidad que se observa es el trabajador, por lo que se descarta a los individuos que no están ocupados o que no reciben pago por su trabajo. Las características de los trabajadores se clasifican en variables de industria y características individuales, además de variables de control de tiempo y por zona rural.

Tabla 2. Variables a utilizar en la estimación

Variables	
Variable Dependiente	Log del ingreso por hora (real)
Características del trabajador y de control	Trimestre-Año Zona Rural/Urbana (dummy) Experiencia Experiencia al cuadrado Años de Educación Casado (dummy) Autoempleado (dummy) Seguridad social (dummy)
Industria	Se incluyen variables dummy de acuerdo para la industria en la que trabaja el individuo. (20 categorías, comercio al menudeo es la base)

Fuente: Elaboración Propia

Zona Rural / Urbana: Esta variable pretenda capturar el efecto de los diferenciales salariales que existen entre zonas urbanas y rurales debido a la mayor infraestructura que existe en las primeras y mejor calidad de servicios. La variable toma el valor de 1 si el individuo vive en una localidad con cien mil habitantes o más, y 0 de otro modo.

[15] Keane (1991) señala la importancia de tener una base que abarque una cantidad significativa de tiempo para poder encontrar evidencia significativa sobre la existencia de salarios de eficiencia.

Experiencia: Debido a que la encuesta no cuenta con información sobre la duración en trabajos anteriores, o la edad en la que el individuo se incorpora al mercado laboral, la medida de experiencia se calcula por medio de la siguiente fórmula:

$$Experiencia = Edad - AñosDeEducación - 6$$

La experiencia potencial nos dice el tiempo total que han tenido oportunidad de trabajar independientemente de si lo hayan hecho o no por lo que podría subestimar aquellos que trabajan y estudian o sobreestimar a aquellos que dejaron de trabajar algún tiempo y luego se reincorporan al mercado laboral. A pesar de las limitaciones de esta variable, es una buena aproximación que sirve para los efectos de esta investigación[16].

Años de Educación: Es una variable de capital humano observable que está directamente relacionada con el salario. Los años de educación se calcularon usando las respuestas a la pregunta del último nivel educativo estudiado y el último año aprobado. Puesto que en la mayor parte del país las preparatorias públicas y privadas tienen una duración de tres años se considera este tiempo para el cálculo de los años de educación completados.

Tabla 3. Variable años de educación.

Años de educación	Nivel educativo
0	No cursó ningún nivel educativo
1	Cursó hasta preescolar
años aprobados	Cursó hasta primaria
6 + años aprobados	Cursó hasta secundaria
9 + años aprobados	Cursó hasta preparatoria
12 + años aprobados	Cursó hasta normal
12 + años aprobados	Cursó carrera técnica
12 + años aprobados	Cursó profesional
16.5 + años aprobados	Cursó maestría
18.5 + años aprobados	Cursó doctorado

Fuente: Elaboración Propia

Casado: Para determinar el valor de esta variable, se da el valor de 1 si la persona está casada o vive en unión libre y 0 de otra forma.

[16] Funciona como variable de control, no se pretende estimar coeficientes insesgados para dicha variable.

Autoempleado: Esta variable toma el valor de 1 si el individuo responde que no tiene jefe y que además tiene un negocio o realiza alguna actividad productiva por su cuenta.

Seguridad Social: La variable dicotómica seguridad social toma el valor de 1 si el individuo respondió que contaba asistencia médica por parte de su trabajo y toma el valor de 0 de otra forma[17]. Esta respuesta incorpora IMSS, hospital o clínica naval, militar o de PEMEX, ISSSTE, ISSSTE estatal u otra institución médica. Esta variable pretende ser una proxy para la formalidad.

Tabla 4. Variables dicotómicas de industria. Clasificación.

Variable	Industria
Ind_1	Agricultura, ganadería, aprovechamiento forestal, pesca y caza
Ind_2	Minería
Ind_3	Electricidad, agua y suministro de gas por ductos al consumidor final
Ind_4	Construcción
Ind_5	Industrias Manufactureras
Ind_6	Comercio al por mayor
Ind_7	Comercio al por menor
Ind_8	Transportes, correos y almacenamiento
Ind_9	Información en medios masivos
Ind_10	Servicios Financieros y Seguros
Ind_11	Servicios inmobiliarios y de alquiler de bienes muebles e intangibles
Ind_12	Servicios profesionales, científicos y técnicos
Ind_13	Dirección de Corporativos y Empresas
Ind_14	Servicios de apoyo a los negocios y manejo de desechos y servicio de remediación
Ind_15	Servicios Educativos
Ind_16	Servicios de Salud y Asistencia Social
Ind_17	Servicios de Esparcimiento, Culturales y deportivos, y otros servicios recreativos
Ind_18	Servicios de alojamiento temporal y de preparación de alimentos y bebidas
Ind_19	Otros servicios excepto actividades del gobierno
Ind_20	Actividades del gobierno y de organismos internacionales y extraterritoriales

Fuente: INEGI 2011

En la tabla 4 se muestra la clasificación de INEGI para las industrias en México. Esta clasificación consta de 20 categorías a partir de las cuales se generan las variables dicotómicas.

Ingreso por Hora Real: Para medir el salario por hora se realiza la transformación de las horas trabajadas semanales en horas trabajadas al mes. Se toma en cuenta de las horas que se trabajan semanalmente de manera habitual y se multiplica por 4.28[18]. Luego, se transforma el salario

[17] Esta respuesta incorpora IMSS, hospital o clínica naval, militar o de Pemex, ISSSTE, ISSSTE estatal u otra institución médica.
[18] Resultado de la división 30/7, que corresponde al número de semanas dentro de un mes en promedio.

mensual reportado en la encuesta en salario real utilizando un deflactor base a la segunda quincena de diciembre de 2010. Por último, se divide el salario real mensual entre el promedio de horas trabajadas al mes.

$$Salario\ por\ hora\ real = \left(\frac{Salario\ real\ mensual}{Horas\ a\ la\ semana * (\frac{30}{7})} \right)$$

iii. Análisis Descriptivo

En esta sección se describen las variables para ver cuál es su relación con el salario real. En la tabla 5 se muestran los estadísticos descriptivos de las variables de características individuales y de control.

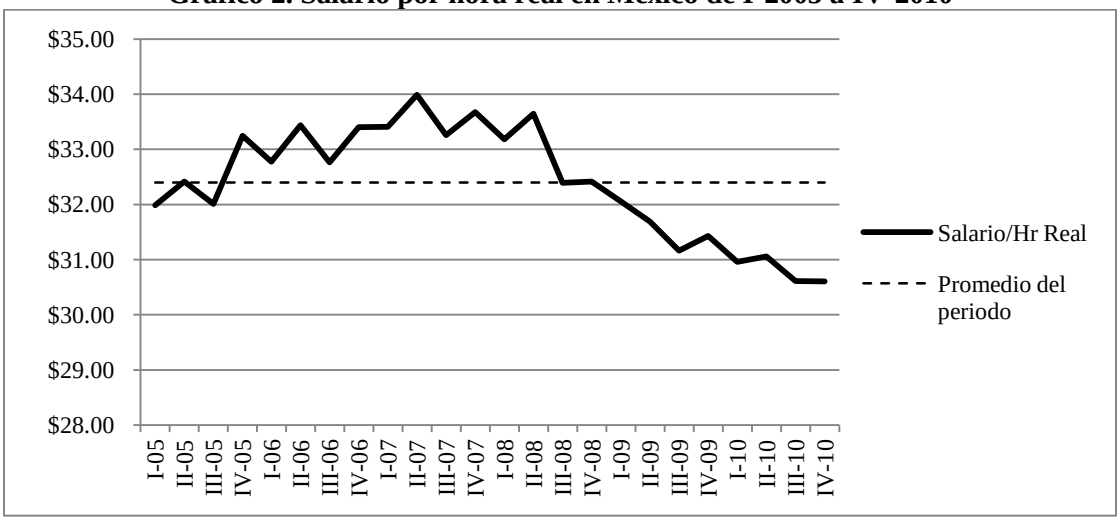
Tabla 5. Estadísticos descriptivos características individuales y variables de control

Variable	Obs.	Media	Desviación Estándar	Min	Max
Salario por hora real	1277155	$32.73142	24.1377	7.970204	166.9805
log (salario por hora real)		3.295264	0.5899564	2.07571	5.117877
Zona Urbana/Rural		0.6242649	0.4843123	0	1
Edad		35.52004	12.45409	14	65
Experiencia		19.87482	13.64059	0	59
Experiencia ^2		581.074	676.2971	0	3481
Años de educación		9.650496	4.513332	0	24.5
Casado		0.6705333	0.4700197	0	1
Seguridad Social		0.4977728	0.4999952	0	1
Autoempleado		0.1822723	0.386069	0	1

Fuente: Elaboración Propia con datos de las ENOE 2005-2010.

En el gráfico 2 se muestra el comportamiento del salario por hora real entre 2005 y 2010. Se puede observar que a inicios de 2005, el salario por hora se fue incrementando hasta que durante el periodo entre el tercer trimestre de 2005 y el segundo trimestre de 2008, el salario por hora se mantuvo por encima del promedio total del periodo. Después, a partir del tercer trimestre de 2008 a raíz de la recesión económica mundial, el salario por hora real se ha venido reduciendo constantemente.

Gráfico 2. Salario por hora real en México de I-2005 a IV-2010



Fuente: Elaboración propia con datos de las ENOE 2005 – 2010.

En la tabla 9 se muestra el salario por hora entre las industrias, la desviación estándar de la media por industria y la participación porcentual de cada industria dentro de la muestra. Las industrias en las que el salario por hora promedio es menor son la industria agrícola (ind_1), hotelera (ind_18), y de servicios de apoyo y manejo de desechos (ind_14), entre otras; mientras que las de mayor salario por hora son las de dirección de corporativos y empresas (ind_13), servicios educativos (ind_15), servicios financieros (ind_10) y la minería principalmente (ind_2).

Las diferencias salariales entre industria se pueden apreciar mejor en la gráfica 3, que muestra la diferencia del salario por hora que se paga en cada industria con respecto a la media nacional de $32.73. La mayor parte de las industrias que pagan salarios por encima de la media tienen una participación muy baja en la muestra, mientras que las industrias con salarios más bajos, como la de construcción, manufactureras o de comercio al por menor tienen las mayores participaciones. Es por esto que parece como si la mayoría de las industrias pagan salarios por encima de la media, pero es porque dichas industrias tienen una participación pequeña del total de los empleados entrevistados.

La información es consistente con lo que se observa en trabajos previos como el de Salas y Zepeda (2003), en el que la industria extractiva (minería) tiene un salario más alto que las demás, mientras que el sector de comercio al por menor y la agricultura suelen tener un salario por debajo

de la media. Sin embargo, estos diferenciales no controlan por ninguna característica observable de los individuos ya que no se toma en cuenta la proporción de trabajadores calificados que varía en cada industria, es decir, que cada industria puede tener diferentes demandas de ocupaciones. Por ejemplo, se puede suponer que la industria manufacturera requiere una mayor proporción de trabajadores fabriles y artesanos que la industria de servicios educativos, la cual requiere mayormente profesionistas y trabajadores de la educación.

Tabla 9. Salario por hora por industria y participación porcentual de la industria en la muestra.

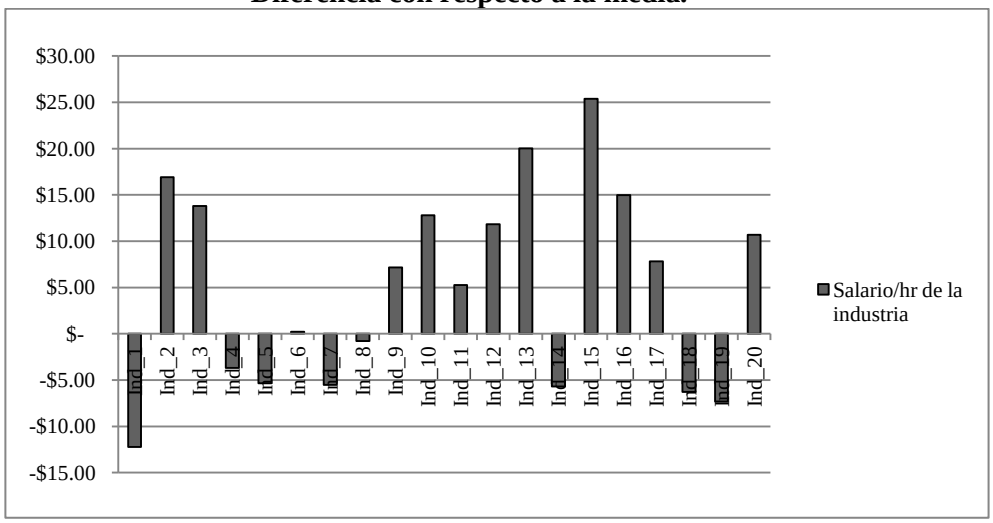
Industria	Salario por hora promedio	Desviación Estándar	Porcentaje Obs.
1 Agricultura, ganadería, aprovechamiento forestal, pesca y caza	$ 20.19	$ 14.71	8.616%
2 Minería	$ 46.82	$ 32.81	0.845%
3 Electricidad, agua y suministro de gas por ductos al consumidor final	$ 46.26	$ 25.57	0.890%
4 Construcción	$ 28.46	$ 16.44	16.617%
5 Industrias Manufactureras	$ 29.20	$ 19.70	19.534%
6 Comercio al por mayor	$ 33.38	$ 23.84	3.277%
7 Comercio al por menor	$ 28.36	$ 20.92	11.553%
8 Transportes, correos y almacenamiento	$ 31.37	$ 21.41	4.234%
9 Información en medios masivos	$ 41.56	$ 29.31	1.019%
10 Servicios Financieros y Seguros	$ 50.43	$ 31.77	0.789%
11 Servicios inmobiliarios y de alquiler de bienes muebles e intangibles	$ 38.98	$ 29.37	0.687%
12 Servicios profesionales, científicos y técnicos	$ 49.16	$ 32.82	2.227%
13 Dirección de Corporativos y Empresas	$ 55.68	$ 37.35	0.041%
14 Servicios de apoyo a los negocios y manejo de desechos y remediación	$ 28.23	$ 20.17	2.080%
15 Servicios Educativos	$ 60.42	$ 31.49	5.140%
16 Servicios de Salud y Asistencia Social	$ 57.11	$ 35.32	2.130%
17 Servicios de Esparcimiento, Culturales y deportivos, y otros servicios recreativos	$ 42.72	$ 32.33	0.985%
18 Servicios de alojamiento temporal y de preparación de alimentos y bebidas	$ 28.24	$ 20.15	5.172%
19 Otros servicios excepto actividades del gobierno	$ 28.41	$ 19.24	6.956%
20 Actividades del gobierno y de organismos internacionales y extraterritoriales	$ 42.72	$ 26.88	7.209%
Total	$ 32.73	$ 24.14	100%

Fuente: Elaboración propia con datos de las ENOE 2005 – 2010.

La hipótesis de categorización nos dice que los individuos se ordenarán de acuerdo a su dotación de capital humano observable (años de educación, experiencia laboral) hacia ocupaciones de salarios altos y hacia las industrias de mayor pago para compensar la demanda de cada industria. Entonces, si existe una demanda diferente de las industrias por capital humano no observable

(habilidad como capacidad de trabajo en equipo, facilidad de palabra, puntualidad, etc.), los individuos con mayores niveles de capital humano no observable se ordenarán hacia las industrias con mayores diferenciales ya que estos reflejan las demandas de cada industria.

Gráfico 3. Salario/hr por industria comparado con el salario/hr promedio total. Diferencia con respecto a la media.



Fuente: Elaboración propia con datos de las ENOE 2005 – 2010. Pesos constantes a precios de diciembre de 2010

El desarrollo de la zona donde se encuentra empleado el individuo también es importante, puesto que existen diferencias significativas entre el salario que se paga en zonas urbanas comparado con las zonas rurales.

Tabla 6. Salario por hora real por zona urbana/rural.

Zona	No. Obs.	Salario/hr	Desv. Est.
Rural	479872	$27.58	0.0296087
Urbana	797283	$35.83	0.0286515

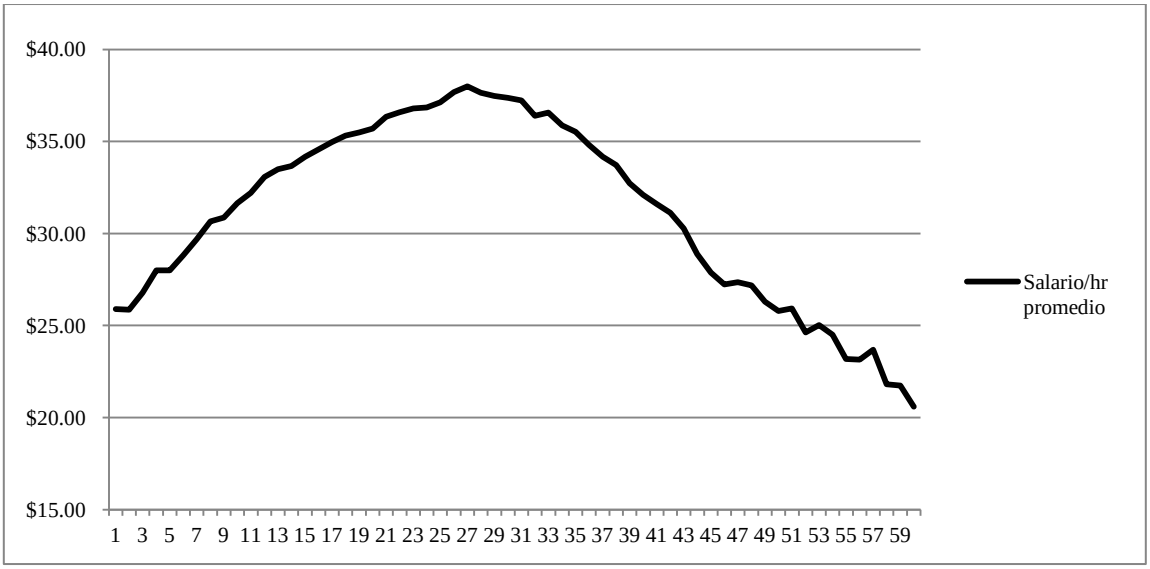
Fuente: Elaboración propia con datos de las ENOE 2005 – 2010.

Esto se muestra en la tabla 6. Se realizó una prueba t para comparar la media del salario por hora real entre los grupos de individuos por zona y se encontró que la diferencia de medias es altamente significativa, lo cual indica que es importante tener este control en la regresión de los determinantes del salario.

La distribución de la experiencia se muestra en el gráfico 4. Esta variable se construye a partir de la variable de edad y de años de educación. Se observa que la relación entre el salario por hora y la experiencia del trabajador no es lineal, sino que es creciente al principio, llegando a un máximo

alrededor de los 26 años de experiencia y después el salario por hora va decreciendo para individuos con mayores niveles de experiencia laboral potencial. Esto nos indica que la relación que existe entre los años de experiencia y el salario por hora real no será lineal, sino que puede modelarse con una forma cuadrática, motivo por el cual en la estimación se incluye la variable años de experiencia al cuadrado.

Gráfico 4. Salario por hora promedio de acuerdo a los años de experiencia.



Fuente: Elaboración propia con datos de las ENOE 2005-2010. Pesos constantes a precios de diciembre de 2010.

Por otro lado, la educación se considera una inversión en capital humano, la cual tiene un impacto en la productividad de las personas y genera rendimientos positivos[19], de modo que, en general, las ocupaciones que reciben mayores pagos son las que requieren más años de educación para ser ejercidas, o en su defecto requieren más años de experiencia.

En la tabla 7 se muestra el salario por hora promedio de acuerdo a los años de educación. Se puede observar una clara tendencia creciente conforme se incrementan los años de educación. Aquellos trabajadores que tienen primaria incompleta o terminada ganan en promedio $24.41 por hora. Este salario se incrementa a $26.20 para trabajadores que tienen estudios de secundaria incompleta o completa. Los individuos que estudiaron hasta preparatoria trunca o la completaron obtienen $31.14 en promedio por cada hora de trabajo, muy cerca pero aún por debajo de la media

[19] Véase (Mincer, 1974)

nacional. Los trabajadores que cuentan con alguna preparación de nivel superior ganan en promedio $50.68 por hora, lo cual representa una diferencia de $19.54 por hora adicional de salario en comparación con aquellos que alcanzaron el nivel de escolaridad medio superior. Esto indica que los individuos con educación superior suelen trabajar en ocupaciones e industrias que pagan salarios más altos. Por último, la brecha salarial más grande se da entre los trabajadores con preparación de posgrado, ya que el salario por hora promedio de este grupo es de $81.36, lo cual representa una brecha de casi $31 pesos por hora más que los trabajadores con estudios de nivel superior y de alrededor de $57 pesos por hora más que los trabajadores con primaria.

Tabla 7. Salario por hora real por años de educación.

Años de educación	Salario/hr promedio
0 a 6 años	$ 24.41
7 a 9	$ 26.20
10 a 12	$ 31.14
13 a 17.5	$ 50.68
18 o más	$ 81.36

Fuente: Elaboración propia con datos de las ENOE 2005 – 2010.

Debido a la forma en que se aplica la encuesta, la variable autoempleado incluye tanto al sector formal como al informal, y como los trabajadores informales en su mayoría no tienen seguridad social, entonces se observa en la tabla 8 que la mayoría de los autoempleados no tiene seguridad social. Por otro lado, gran parte de los empleados no tiene seguridad social de ningún tipo, lo que sugiere un alto nivel de informalidad en la muestra, o que existe incumplimiento de las normas laborales por parte de los patrones para evitar el pago de cuotas.

Tabla 8. Salario por hora para trabajadores autoempleados / con seguridad social

	No tiene seguridad social	Tiene seguridad social	Total
No autoempleado	$ 23.43	$ 37.24	$ 31.80
(No. Observaciones)	411071	633294	1044365
Autoempleado	$ 36.73	$ 52.27	$ 36.89
(No. Observaciones)	230351	2439	232790
Total	$ 28.21	$ 37.30	$ 32.73
(No. Observaciones)	641422	635733	1277155

Fuente: Elaboración propia con datos de las ENOE 2005-2010. Pesos constantes a precios de diciembre de 2010.

También se observan diferencias significativas[20] entre el salario que se paga a individuos que son autoempleados y no autoempleados ya que los primeros ganan en promedio $36.89 por hora mientras que los no autoempleados ganan $31.80. La diferencia es aún más drástica entre los individuos con seguridad social y los que no tienen, ya que los primeros ganan $37.30 por hora mientras que los que no tienen seguridad social ganan en promedio $28.21 por hora, lo que representa una diferencia de casi nueve pesos por hora además de la seguridad social. Por último, aquellos individuos que son autoempleados y tienen seguridad social, es decir, que cuentan con un negocio propio o trabajan por su cuenta y que tienen acceso a servicios de salud por parte de su trabajo ganan en promedio $52.27, lo que representa una diferencia con respecto a la media total de casi veinte pesos, mientras que los que no tienen ninguna seguridad social y que además son empleados ganan en promedio $23.43 por hora, lo que representa en total una diferencia de casi 9 pesos con respecto a la media total y de $28.84 pesos en promedio entre los individuos. Debido a que existen diferencias tan importantes entre los distintos grupos es que se incluye a estas dos variables de control en el análisis de los diferenciales salariales como controles y características del individuo.

En el análisis no se incluyen controles por la ocupación del individuo, es decir, por la posición que ocupa en el trabajo ya que esta información en la encuesta está dividida en clasificaciones que se encuentran altamente correlacionadas con las clasificaciones de la industria. Las clasificaciones de ocupaciones que utiliza el INEGI tienden a estar muy correlacionadas por definición, por ejemplo, la clasificación de trabajadores de la educación, la cual está altamente correlacionada con la industria de servicios educativos, y por lo tanto, la gran mayoría de las observaciones de trabajadores de la educación pertenecen a servicios educativos y viceversa. Debido a esto y para prevenir alta correlación entre las variables explicativas se utilizan controles únicamente por industria.

[20] Se realizaron pruebas t para probar si el salario por hora promedio entre autoempleados y no autoempleados es diferente. También se hizo la prueba para individuos con seguridad social y sin seguridad social.

V. ESTIMACIONES Y RESULTADOS

i. Categorización o 'Sorting'

La primera parte de los resultados consiste en la probar la hipótesis de categorización a través de la estimación por mínimos cuadrados ordinarios de la ecuación (7). En la tabla 9 se muestran los resultados de la estimación de los diferenciales salariales entre industrias. Los errores estándar que aparecen en la tabla ya han sido corregidos por problemas de heteroscedasticidad y todos los coeficientes son significativos al 1%. A primera vista, los resultados sugieren que los diferenciales son muy amplios de una industria a otra y también que la importancia de las características individuales en la determinación del salario es muy grande. Las variables individuales muestran que en promedio, un año más de educación rinde un 5.28% de salario por hora más, mientras que los años de experiencia rinden un 2.33%[21]. También en promedio, las personas casadas tienden a ganar 7.23% más, manteniendo lo demás constante. La variable de autoempleado tiene un coeficiente positivo que implica que los salarios son 23.23% mayores para personas que trabajan por su cuenta o tienen negocios propios que aquellos que tienen un patrón o jefe directo.

Para las variables dicotómicas de la industria, la industria base es comercio al por menor. A excepción de la industria agrícola, todas las industrias tienen coeficientes positivos y significativos, lo que sugiere que la industria que recibe el menor salario después de controlar por variables de individuales, regionales y de tiempo es la industria agrícola seguida por la de comercio al por menor. De acuerdo a la regresión, la industria de servicios educativos es la que paga mayores salarios con un coeficiente que representa un 49.07% más de salario en comparación a los que trabajan en el comercio al por menor, manteniendo todo lo demás constante. Las industrias de minería, electricidad, agua y suministro de gas, servicios de salud y dirección de corporativos y empresas tienen los premios salariales más altos. Por el otro lado, las industrias que pagan menores salarios son el

[21] Rendimiento calculado al valor de experiencia promedio en la muestra, 19.51 años.

comercio al por menor, industria manufacturera, agricultura, y servicios de remediación y manejo de desechos, principalmente, junto con servicios de alojamiento temporal y otros servicios.

Tabla 9. Estimación por Mínimos Cuadrados Ordinarios de los determinantes del Salario. Ecuación 7.

Variable Dependiente: log(salario por hora)	Coeficiente	Error Estándar	t	P>t
Años de Educación	0.0528428*	0.0001384	381.72	0.0000
Experiencia	0.0159021*	0.0001162	136.85	0.0000
Experiencia^2	-0.0001915*	2.22E-06	-86.09	0.0000
Casado	0.0723581*	0.0010546	68.61	0.0000
Autoempleado	0.232377*	0.0014663	158.48	0.0000
Seguridad Social	0.1417204*	0.0010485	135.17	0.0000
Agricultura, ganadería, aprovechamiento forestal, pesca y caza	-0.017673*	0.0020956	-8.43	0.0000
Minería	0.4313788*	0.0054261	79.5	0.0000
Electricidad, agua y suministro de gas por ductos al consumidor final	0.3843415*	0.0045781	83.95	0.0000
Construcción	0.2240986*	0.0016352	137.04	0.0000
Industrias Manufactureras	0.0911264*	0.0015939	57.17	0.0000
Comercio al por mayor	0.1309962*	0.0026692	49.08	0.0000
Transportes, correos y almacenamiento	0.1080402*	0.0025709	42.02	0.0000
Información en medios masivos	0.2257577*	0.0049352	45.74	0.0000
Servicios Financieros y Seguros	0.3615740*	0.0054125	66.8	0.0000
Servicios inmobiliarios y de alquiler de bienes muebles e intangibles	0.2497664*	0.0060306	41.42	0.0000
Servicios profesionales, científicos y técnicos	0.2656682*	0.0035878	74.05	0.0000
Dirección de Corporativos y Empresas	0.4247136*	0.0256409	16.56	0.0000
Servicios de apoyo a los negocios y manejo de desechos y de remediación	0.0160125*	0.0032721	4.89	0.0000
Servicios educativos	0.4907250*	0.0024399	201.12	0.0000
Servicios de Salud y Asistencia Social	0.3947573*	0.0034931	113.01	0.0000
Servicios de Esparcimiento, Culturales y deportivos, y otros servicios recreativos	0.3753572*	0.0060939	61.6	0.0000
Servicios de alojamiento temporal y de preparación de alimentos y bebidas	0.0576509*	0.0024036	23.99	0.0000
Otros servicios excepto actividades del gobierno	0.0600676*	0.002124	28.28	0.0000
Actividades del gobierno y de organismos internacionales y extraterritoriales	0.2656680*	0.0020832	127.53	0.0000
Zona Urbana	0.1028876*	0.0009451	108.87	0.0000
Constante	2.1528040*	0.0030019	717.14	0.0000
Observaciones: 1277155	**R-cuadrada: 0.3659**	**Varianza del error: 0.22069**		
***Significativo al 1%**	****Significativo al 5%**			

Nota: La regresión incluye variables dummy por trimestre/año. Fuente: Elaboración propia.

Aunque estos resultados parecen sugerir la existencia de verdaderos diferenciales salariales entre industrias, se quiere probar que dichos diferenciales están sesgados debido a que la habilidad no observada no es tomada en cuenta por estas estimaciones. Sin embargo, es posible observar otras

características de capital humano que están correlacionados con la industria y la habilidad, es decir, que las industrias tienen diferentes demandas por habilidad no observada y que por lo tanto aquellas que requieren mayor capacidad lo compensan con mayores salarios. Para probar esta aseveración se estiman las dos regresiones auxiliares (8) y (9) que se muestran en las tablas 10 y 11 respectivamente.

Tabla 10. Estimación por Mínimos Cuadrados Ordinarios de la regresión auxiliar 1. Ecuación 8.

Variable Dependiente: log(salario por hora)	Coeficiente	Error Estándar	t	P>t
Años de educación	0.0656884*	0.0001233	532.8100	0.0000
Experiencia	0.0191844*	0.00012	159.8300	0.0000
Experiencia^2	-0.0002183*	2.31E-06	-94.5200	0.0000
Casado	0.0647793*	0.0010996	58.9100	0.0000
Autoempleado	0.2046567*	0.0014943	136.9600	0.0000
Seguridad Social	0.1821523*	0.0009954	183.0000	0.0000
Constante	2.235489*	0.001607	1391.0600	0.0000
Observaciones: 1277155	**R-cuadrada: 0.3103**		**Varianza del Error: 0.24003**	
***Significativo al 1%**	****Significativo al 5%**			

Fuente: Elaboración propia.

Para poder entender un poco mejor los coeficientes desde el enfoque de la categorización, es necesario examinar los resultados de la tabla 10. La ecuación (8) es simplemente el efecto que tienen las características observables del individuo sobre el logaritmo de su salario por hora y la relación que hay entre un mayor pago un mayor capital humano independientemente de la zona, industria o trimestre que se analice. Los resultados muestran que el salario por hora se incrementa 6.56% al aumentar la educación en un año, independientemente de la industria en que se encuentre empleado. Por otro lado, un año adicional de experiencia incrementa el salario en un 1.03% en promedio, tomando en cuenta el coeficiente de experiencia al cuadrado y calculado con el valor medio de la experiencia de los trabajadores en la muestra.

De este primer resultado, se puede observar que existe una relación positiva entre las variables de años de educación y el salario, y lo mismo sucede con los años de experiencia. Es decir, que sí hay una compensación salarial al capital humano observable, y que es altamente significativo independientemente de la industria en la que se utilice. Para probar si es que existen diferentes demandas por capital humano entre las industrias se estima la ecuación (9). En esta regresión la

variable dependiente es el efecto salarial total estimado en la tabla 9 para cada individuo y se corre contra las características observables de ese trabajador.

Tabla 11. Estimación por Mínimos Cuadrados Ordinarios de la regresión auxiliar 2. Ecuación 9.

Variable Dependiente: Efecto Total	Coeficiente	Error Estándar	T	P>t
Años de educación	0.0128456*	0.0000345	372.3500	0.0000
Experiencia	0.0032823*	0.0000339	96.7000	0.0000
Experiencia^2	-0.0000268*	6.49E-07	-41.2700	0.0000
Casado	-0.0075788*	0.0003148	-24.0700	0.0000
Autoempleado	-0.0277203*	0.0003651	-75.9200	0.0000
Seguridad Social	0.0404319*	0.0002913	138.8100	0.0000
Constante	0.0826848*	0.0004607	179.4900	0.0000
Observaciones: 1277155	**R-cuadrada: 0.1773**		**Varianza del Error: 0.01935**	
***Significativo al 1%**	****Significativo al 5%**			

Fuente: Elaboración propia.

Para la interpretación de los resultados de la estimación de la ecuación (9), se toman los diferenciales de la industria, zona y trimestre (el efecto salarial total) y se busca ver si existe una correlación entre las características observables del trabajador el premio salarial de cada industria. En otras palabras, si existe dicha correlación, entonces existe una demanda que difiere entre industrias por las características de capital humano del trabajador. Por lo tanto, si la hipótesis de categorización (*'sorting'*) es correcta, entonces las características del individuo que generan mayores salarios también tienen mayor poder explicativo sobre los diferenciales de las industrias que pagan más, es decir, los coeficientes de las variables independientes deben de tener el mismo signo.

El coeficiente de la educación, por ejemplo, indica que un año más de educación genera un 6.56% más de salario y además, categoriza a los individuos en industrias/ocupaciones que pagan un 1.28% más. En el caso de la experiencia, la media es de 19.51 años, lo cual en promedio genera un salario 1.06% mayor al de un trabajador sin experiencia, independientemente de la industria,

ocupación u otras características. Además, una persona con 19.51 años de experiencia tendería a trabajar en una industria que paga 0.22% más[22].

Contrario a lo esperado de acuerdo a la hipótesis de categorización, los hombres casados ganan un 6.47% más que los solteros, sin embargo los resultados de la tabla 11 sugieren que tienden a trabajar en industrias que pagan 0.757% menos. En el caso de los autoempleados, ellos ganan en promedio 20.46% más que los empleados manteniendo lo demás constante, no obstante tienden a trabajar en industrias que pagan 2.77% menos. Para los fines de contrastar la hipótesis de categorización, estos resultados indican que la categorización existe únicamente a través de las variables de capital humano como los años de educación pero no a través de características individuales como lo son el estado civil o el tipo de empleo, ya que si así fuera, los coeficientes de estas variables tendrían el mismo signo tanto en la tabla 10 como en la tabla 11.

Con base en estos resultados, se encuentra evidencia de que la hipótesis de categorización por medio de capital humano observable es correcta. Es decir, los trabajadores con mayor capital humano tienden a ganar más porque se autoseleccionan hacia industrias que tienen mayores demandas y por lo tanto dan mayores pagos a ésas características.

Como conclusión de esta primera estimación, se puede decir que los individuos sí se categorizan y filtran entre las industrias por medio de sus características observables, principal y más drásticamente por medio de los años de educación. De modo que si los individuos se categorizan en las diferentes industrias por medio de su capital humano (incluyendo el no observable como habilidad), entonces los diferenciales salariales entre industrias no podrían ser evidencia concluyente de salarios de eficiencia, y sería necesario eliminar aquellas características de capital humano no observable que afectan la productividad del trabajador. A continuación, se estima la ecuación de los determinantes del salario utilizando una metodología de efectos fijos para

[22] La inclusión de la experiencia al cuadrado y que su coeficiente sea negativo y significativo, sugiere que hay rendimientos decrecientes en los años de experiencia, por lo que la interpretación de los años de experiencia debe de tomar en cuenta esto.

intentar controlar por aquellas características no observables como la habilidad que se comportan como un efecto fijo en el tiempo e individual de cada trabajador.

ii. Efectos Fijos

Esta metodología asume que la habilidad del individuo es un efecto fijo que no varía en el tiempo, entonces al transformar la ecuación (7) en la ecuación (12), dicho efecto es eliminado de la regresión. La estimación de los diferenciales salariales por medio de efectos fijos se muestra en la tabla 12.

Esta regresión consta de 563,498 trabajadores, con un total de 1,277,155 observaciones en un panel no balanceado. Se estimó la regresión (7) por efectos fijos, es decir, la transformación de efectos fijos (ecuación 12) por MCO. Se incluyeron como variables explicativas los años de educación, experiencia, experiencia al cuadrado, entre otras características individuales y se incluyeron las variables dicotómicas de industria para verificar si los coeficientes estimados bajo esta metodología cambiaban con respecto a los estimados en la tabla 9. No se incluyeron los controles por trimestre puesto que no tienen relevancia para el análisis, además de que dicho efecto sería el efecto promedio de los trimestres.

Los resultados muestran que los diferenciales salariales entre industrias en realidad no son tan grandes como se mostraba en la primera estimación por MCO. Únicamente en el caso de la industria agrícola y servicios de apoyo y manejo de desechos, los diferenciales salariales se incrementaron (en valor absoluto) con respecto al comercio al por menor, al pasar de ser -1.76% a -5.28% y de 1.6% a 3.23%. En todas las demás industrias, los diferenciales salariales se redujeron en comparación a la estimación de corte transversal. En promedio, los diferenciales salariales se redujeron un 35% con respecto a los obtenidos en la estimación de la ecuación (7).Este resultado implica que aquellas características de los individuos como su nivel de habilidad tienen un efecto importante en la determinación salarial, y en parte explican los diferenciales salariales ya que la disminución de magnitud es muy significativa.

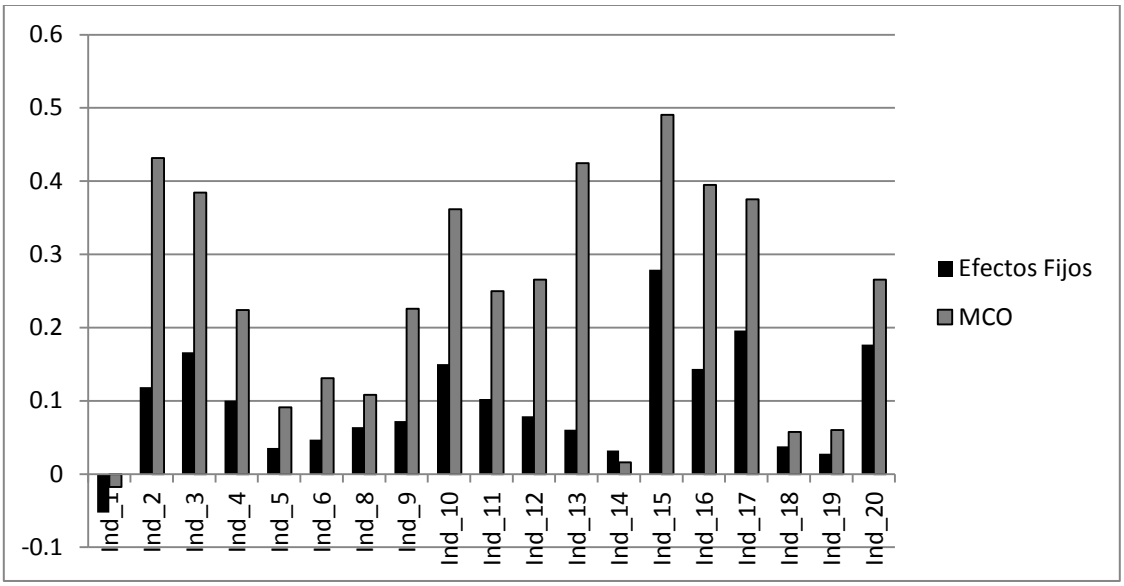
Tabla 12. Estimación por Efectos Fijos de los determinantes del salario. Ecuación 12.

Variable Dependiente: log(salario por hora)	Coeficiente	Error Estándar	t	P>t
Años de Educación	0.0897877*	0.0250717	3.5800	0.0000
Experiencia	0.0118149*	0.0015168	7.7900	0.0000
Experiencia^2	-0.0002983*	0.0000324	-9.2100	0.0000
Casado	-0.0716813	0.0912236	-0.7900	0.4320
Autoempleado	0.1626734*	0.0024411	66.6400	0.0000
Seguridad Social	0.0632879*	0.0017023	37.1800	0.0000
Agricultura, ganadería, aprovechamiento forestal, pesca y caza	-0.0528052*	0.0047291	-11.1700	0.0000
Minería	0.1187739*	0.0096363	12.3300	0.0000
Electricidad, agua y suministro de gas por ductos al consumidor final	0.1664075*	0.0106532	15.6200	0.0000
Construcción	0.1004131*	0.0035766	28.0800	0.0000
Industrias Manufactureras	0.0356509*	0.0032453	10.9900	0.0000
Comercio al por mayor	0.0471181*	0.0035476	13.2800	0.0000
Transportes, correos y almacenamiento	0.0642022*	0.0056164	11.4300	0.0000
Información en medios masivos	0.0721969*	0.008848	8.1600	0.0000
Servicios Financieros y Seguros	0.1500091*	0.0125302	11.9700	0.0000
Servicios inmobiliarios y de alquiler de bienes muebles e intangibles	0.1026738*	0.0085191	12.0500	0.0000
Servicios profesionales, científicos y técnicos	0.0789114*	0.0062075	12.7100	0.0000
Dirección de Corporativos y Empresas	0.0607385**	0.0252828	2.4000	0.0160
Servicios de apoyo a los negocios y manejo de desechos y de remediación	0.0323643*	0.0050163	6.4500	0.0000
Servicios Educativos	0.2791588*	0.0087152	32.0300	0.0000
Servicios de Salud y Asistencia Social	0.1437757*	0.0089233	16.1100	0.0000
Servicios de Esparcimiento, Culturales y deportivos, y otros servicios recreativos	0.1959031*	0.0107502	18.2200	0.0000
Servicios de alojamiento temporal y de preparación de alimentos y bebidas	0.0376744*	0.0057792	6.5200	0.0000
Otros servicios excepto actividades del gobierno	0.0277961*	0.0039661	7.0100	0.0000
Actividades del gobierno y de organismos internacionales y extraterritoriales	0.1770033*	0.0055873	31.6800	0.0000
Zona Urbana	-0.0130717	0.0218979	-0.6000	0.5510
Constante	2.295426*	0.2538533	9.0400	0.0000
Observaciones: 1277155 Grupos: 562398 R-cuadrada: 0.2444 Varianza del Error: 0.33309				
Varianza del Efecto fijo: 0.51181 Valor Medio del Efecto Fijo: 2.295426				
***Significativo al 1% **Significativo al 5%**				

Fuente: Elaboración propia.

En el gráfico 5 se puede observar el cambio de en la magnitud de los diferenciales entre las estimaciones de corte transversal y de efectos fijos por individuo. De acuerdo a la teoría, las disminuciones más importantes en los diferenciales son en aquellas industrias en las que se valora más la habilidad de los individuos y otras características de capital humano no observables.

Gráfico 5. Comparación de diferenciales salariales entre industrias de MCO y Efectos Fijos.



Fuente: Elaboración Propia

Por ejemplo, en la industria 13, que corresponde a la dirección de corporativos y empresas, la reducción del diferencial salarial fue de un 85.7%, lo que quiere decir que una vez que se controla por aquella habilidad gerencial no observada y única del trabajador, el coeficiente que se observa a primera vista en realidad refleja una compensación monetaria por sus capacidades productivas. De igual manera, en la industria 2, correspondiente a la minería, aquellas características individuales no capturadas que son demandadas por la industria se eliminan del error al usar efectos fijos, reduciendo el diferencial salarial de 43.13% más que el comercio al por menor a sólo 11.87% más.

Cabe resaltar que siguen existiendo los diferenciales salariales, y que todos son altamente significativos, lo que podría representar evidencia de la existencia de verdaderos diferenciales salariales. Por lo tanto, esta estimación aporta evidencia que apoya el resultado de la metodología anterior al mostrar que buena parte de los diferenciales se elimina al controlar por efectos fijos. Sin embargo, al seguir existiendo los diferenciales se puede considerar como evidencia de que en algunas industrias se pagan salarios de eficiencia, aunque para poder concluir esto sería necesario un estudio más profundo que tome en cuenta también características de la ocupación y del trabajo.

Es decir, los remanentes de los diferenciales salariales para las industrias que se observan en la estimación de efectos fijos podrían pero no necesariamente deben ser considerados evidencia de que algunas industrias pagan salarios de eficiencia, puesto que bien podrían ser atribuidos a efectos de ajuste del mercado laboral o a características no observadas del trabajo del individuo.

Hasta el momento se ha encontrado evidencia de categorización o *'sorting'* de los individuos a través de sus variables de capital humano que explican parte de los diferenciales puesto que existen diferentes demandas por parte de cada industria y que al controlar también por características no observables como su habilidad, los diferenciales se reducen drásticamente, sugiriendo que dichas características explican en buena parte las brechas salariales entre industrias.

i. Medida de sesgo

Aunque se ha encontrado evidencia que apoya la teoría neoclásica, esta evidencia no es concluyente puesto que los diferenciales siguen estando ahí, sin embargo, el objeto de este estudio es mostrar la importancia de las características no observables en la estimación debido a la correlación que existe entre el capital humano no observable (la habilidad) y los altos salarios en algunas industrias.

Para estimar el sesgo por omisión de la habilidad en los coeficientes, se asumirá que la teoría neoclásica es correcta y que los diferenciales salariales son iguales a cero, pero que existe un efecto fijo de la habilidad del individuo, de modo que la ecuación de los determinantes del salario (7) se transforma en la ecuación (13).

Utilizando la metodología de Murphy y Topel (1990) se estimará una medida del sesgo que causa la omisión del efecto fijo en la estimación de la ecuación de corte transversal, bajo el supuesto de que los efectos de la industria y ocupación son iguales a cero. Partiendo de las ecuaciones (13) y (14) se obtiene la fórmula de sesgo (15) para variables omitidas:

$$SESGO = \varphi\sigma_{\alpha}^{2}/\sigma_{\varepsilon}^{2}$$

Donde el coeficiente φ representa la correlación entre la habilidad no observada y el premio salarial de la industria, σ_{ε}^{2} es la varianza del error en la estimación de los determinantes del salario dado que no se incluye la habilidad y σ_{α}^{2} es la varianza de la habilidad. El verdadero valor de estos parámetros es desconocido por lo que haciendo uso de los parámetros estimados en las regresiones anteriores se obtendrá una medida aproximada del sesgo.

Puesto que se hace el supuesto de que los diferenciales salariales de las industrias son cero, σ_{ε}^{2} se estimará con el valor de la varianza del error de la regresión auxiliar 1 (ecuación 8), que es 0.24003. Como estimador de σ_{α}^{2} se utiliza la varianza estimada de la habilidad no observada (α_i) obtenida en la ecuación de efectos fijos de 0.51181. Y por último, el coeficiente φ que es el "efecto de categorización" o la correlación entre la habilidad y el premio salarial se estimará a través del porcentaje de ganancia salarial que obtiene un trabajador al desplazarse a una industria de mayor pago (coeficiente β en ecuación 8) entre el porcentaje de ganancia salarial en general (coeficiente β en ecuación 9). En este caso, los individuos se categorizan a través de la educación, entonces aquellos con un año más de educación ganan 6.56% más y además trabajan en industrias que pagan en promedio 1.28% más, por lo que el efecto de categorización es igual a 0.1951= (1.28/6.56). De este modo, si los individuos se categorizan a través de su habilidad de la misma manera que lo hacen con los años de educación (es decir, misma correlación) el sesgo en la estimación de MCO sería del 41.6% = [(0.1951*0.51181)/0.24003].

Esto significa que los diferenciales salariales que se observan en la primera estimación de la ecuación (7) están sesgados hacia arriba en un 41.6% debido a que los individuos se categorizan hacia industrias de mayor demanda a través de su habilidad no observada y esto les ofrece una compensación salarial. Estos resultados son similares al 35% de reducción de los diferenciales al controlar por efectos fijos, lo que podría interpretarse como que el efecto de categorización de la habilidad es muy similar al efecto de categorización de los años de educación.

Sin embargo, el verdadero efecto de categorización a través de la habilidad no observada de los individuos no es conocido, es decir, la magnitud que tiene para desplazar a los trabajadores más

hábiles hacia industrias de mayor pago es desconocida. En la tabla 13 se muestra una tabulación del efecto de sesgo que puede existir en la estimación de los diferenciales salariales si no se considera a la habilidad dentro de la regresión.

Tabla 13. Estimación del Sesgo por Omisión de la Habilidad No Observada

Efecto Categorización φ	SESGO	
0.05	10.66%	
0.1	21.32%	$\widehat{\sigma_{\alpha}^{2}}$
0.15	31.98%	0.51181
0.2	42.65%	
0.25	53.31%	
0.3	63.97%	$\widehat{\sigma_{\varepsilon}^{2}}$
0.35	74.63%	0.24003
0.4	85.29%	
0.45	95.95%	
0.5	106.61%	

Fuente: Elaboración Propia

En la tabla se puede observar que a medida que el efecto de categorización es mayor, mayor es el sesgo en la estimación de corte transversal. Ahora bien, que si la reducción en los diferenciales del 35% que se obtuvo en la regresión de efectos fijos es correcta, entonces esto indicaría que el efecto de categorización a través de la habilidad debe encontrarse entre 0.15 y 0.20, lo que quiere decir que la razón de las ganancias por desplazamiento a industrias de mayor pago con respecto a las ganancias que genera la habilidad se encuentra en ese rango. Por lo tanto, se puede concluir que si las características de capital humano no observables categorizan a los individuos de la misma forma en que lo hacen con sus características de capital humano observable, entonces la medida de sesgo puede llegar a ser muy grande, de tal modo que la categorización de los individuos a través de sus características no observables como la habilidad explique entre el 35% y 42% de los diferenciales salariales que se observan en una estimación por MCO.

VI. CONCLUSIONES.

En la estimación de los diferenciales salariales entre industrias se encontró evidencia de categorización o '*sorting'* de los individuos más hábiles hacia las industrias de mayores salarios, y que la omisión de dicho efecto resultará en estimadores de los diferenciales sesgados. Consistente con la teoría de capital humano, los individuos más preparados y con más experiencia tienden a ganar salarios mayores que los demás. En promedio se observó que un año más de educación genera un rendimiento de 8.97% más de salario por hora, mientras que la experiencia genera un rendimiento de 1.12% más de salario por hora alrededor del valor medio. Además también se observó que las industrias tienen demandas diferentes por capital humano al encontrar una correlación importante entre los años de educación y experiencia y el premio salarial de las industrias. A mayores valores de educación y experiencia, mayor el premio salarial, no así con variables de características individuales como el estado civil o el ser autoempleado.

Después al utilizar el modelo de efectos fijos se eliminó el sesgo causado por la omisión de las características no observadas de los individuos y a su vez, se obtuvieron mejores estimadores de los diferenciales salariales, reduciéndose estos en un 35% en promedio, pero aún así persistiendo las brechas con coeficientes altamente significativos. Las industrias en las que la correlación entre variables de capital humano y el premio salarial es mayor fueron las industrias en las que la brecha disminuyó más drásticamente, como la de dirección de corporativos que pasó de tener un rendimiento de 42.47% mayor que el comercio al por menor a tan solo 6.07% mayor. Las industrias agrícola y de servicios de apoyo y manejo de desechos incrementaron su diferencial, posiblemente por la correlación contraria entre el capital humano y el premio salarial, ya que son industrias que tienen muy poca participación de trabajadores calificados y requieren menos años de educación.

En la tercera parte de la metodología se estimó una medida del sesgo como la que realizan Murphy y Topel (1990) y se encontraron resultados similares[23]. Se encontró que el sesgo puede estar entre 35% y 42% al no incluir la habilidad de los individuos dependiendo del verdadero valor del efecto de categorización, que se refiere a la correlación que existe entre la habilidad y el premio salarial de cada industria.

La evidencia sugiere que la heterogeneidad de los individuos tiene una gran correlación con los premios salariales, es decir, que los empleadores sí compensan a sus trabajadores por aquellas características que los hacen productivos y no necesariamente por salarios de eficiencia. Aunque es posible que las brechas que persisten después de controlar por efectos fijos sean evidencia de salarios de eficiencia ya que existen problemas de asimetría de la información en los mercados que pueden dar lugar a salarios más altos en el equilibrio, la evidencia econométrica no es suficiente para concluir esto, puesto que también podrían ser atribuibles estas brechas a otros factores de control que no se incluyen en la regresión, como bien podrían ser características de la ocupación o participación en la industria de trabajadores calificados.

Este trabajo pretende hacer una aportación a la investigación sobre los diferenciales salariales entre industrias para México para conocer mejor el mercado laboral mexicano y desarrollar mejores metodologías o mejorar las bases de datos de modo que se obtengan resultados más concluyentes que permitan entender qué otras características mantienen los diferenciales que a su vez llevan a tasas de desempleo. Sin embargo, algunas de las limitaciones de esta investigación se encuentran en la definición de las clasificaciones de las industrias y de las ocupaciones en la base de datos, ya que la inclusión de controles por ocupación o características del trabajo permitirían determinar más precisamente los diferenciales salariales. De igual manera se podría extender este análisis para controlar por variables que no se pudieron incluir en este estudio: como la afiliación sindical, historia laboral, etc.

[23] Murphy y Topel (1990) estiman que la habilidad no observada explica alrededor del 30% de los diferenciales salariales de corte transversal.

VII. BIBLIOGRAFIA

Alarcón, D., & McKinley, T. (1997). The rising contribution of labor income to inequality in Mexico. *The North American Journal of Economics and Finance* , 201-212.

Blackburn, M., & Neumark, D. (1991). Unobserved ability, efficiency wages, and the inter-industry wage differentials. *Quarterly journal of economics* .

Combes, P., Duranton, G., & Gobillon, L. (2005). Spatial wage disparities: sorting matters! *GREQAM*.

Cortez, W. (2005). Dispersión y Estabilidad de las diferencias salariales interestatales en México, 1984-2000. *Investigación Económica. UNAM* , 123-158.

Dickens, W. T., & Katz, L. F. (1987). Inter-Industry Wage Differences and Industry Characteristics. In L. K, & L. J. S., *Unemployment and the Structure of Labor Markets* (pp. 28-89). New York: Basil Blackwell.

Gibbons, R., & Katz, L. (1991). Job mobility and inter-industry wage differences. *(NBER, mimeo)* .

INEGI. (2007). Conociendo la base de datos de la ENOE. México.

INEGI. (2009). Estadísticas de la dinámica laboral en México 2005-2007. México.

Katz, L. (1986). Efficiency wage theories: a partial evaluation. *NBER Macroeconomics Annual I* .

Katz, L., & Summers, L. (1989). Industry Rents: Evidence and Implications. *Brookings Papers on Economic Activity: Microeconomics* .

Keane, M. P. (1991). Individual heterogeneity and interindustry wage differentials. *Federal Reserve Bank of Minneapolis and University of Minnesota* .

Krueger, A. a. (1988). Efficiency wages and the inter-industry wage structure. *Econometrica, 56* .

Krueger, A., & Summers, L. (1987). Reflections on the inter-industry wage structure. In K. Lang, & L. J. (eds.), *Unemployment and the structure of labor markets.* Oxford: Basil Blackwell.

Meza González, L. (2003). Apertura comercial y cambio tecnológico. Efectos en el mercado laboral mexicano. *Fondo de Cultura Económica. vol. LXX* , 457-507.

Mincer, J. (1974). *Schooling, Experience and Earnings. Human Behavior and Social Institutions.* New York: National Bureau of Economic Research.

Murphy, K. M., & Topel, R. H. (1989). Ability biases in models of earnings: new methods and evidence. *(University of Chicago, mimeo)* .

Murphy, K. M., & Topel, R. H. (1990). Efficiency wages reconsidered: Theory and Evidence. In Y. Weiss, & G. (. Fishelson, *Advances in the theory and measurement of unemployment.* London: Macmillan.

Murphy, K. M., & Topel, R. H. (1987). Unemployment, risk and earnings. In K. Lang, & J. (. Leonard, *Unemployment and the structure of labor markets.* Oxford: Basil Blackwell.

Salas, C., & Zepeda, E. (2003). Employment and Wages in Contemporary Mexico. In *State of working in México.* The solidary center.

Salas, I., & García, F. (2010). Salarios de eficiencia y el desempleo en México. *Universidad de Guadalajara* .

Salop, S. (1979). A model of the natural rate of unemployment. *American Economic Review* .

Shapiro, C., & Stiglitz, J. (1982). Equilibrium Unemployment as a worker discipline device. *mimeo., Princenton University* .

Stiglitz, J. (1976). prices and queues as screening devices in competitive markets. *IMSSS technical report no. 212 Stanford University* .

Stiglitz, J. (1974). Wage determination and Unemployment in L.D.C.'s: The labour turnover model. *Quarterly Journal of Economics* .

Weiss, A. (1980). Job Queues and layoffs in labor markets with flexible wages. *Journal of Political Economy* .

VIII. APÉNDICE.

i. Demostración 1

El modelo a estimar es el siguiente:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + e_i \qquad \text{(A1)}$$

Pero suponga que no se incluye a la variable x_2 y se estima:

$$y_i = \widetilde{\beta_0} + \widetilde{\beta_1} x_{1i} + u_i \qquad \text{(A2)}$$

La omisión de x_2 causará que la información que aporte esta variable se vaya al error, de modo que $u_i = \beta_2 x_{2i} + e_i$. Entonces, al calcular el coeficiente $\widetilde{\beta_1}$ por mínimos cuadrados ordinarios:

$$\widetilde{\beta_1} = \frac{\sum_{i=1}^{n}(x_{i1} - \bar{x}_1)\, y_i}{\sum_{i=1}^{n}(x_{i1} - \bar{x}_1)^2} \qquad \text{(A3)}$$

Sustituyendo el valor de $y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + e_i$, se obtiene:

$$\widetilde{\beta_1} = \frac{\sum_{i=1}^{n}(x_{i1} - \bar{x}_1)\,(\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + e_i)}{\sum_{i=1}^{n}(x_{i1} - \bar{x}_1)^2}$$

Entonces, el valor esperado de $\widetilde{\beta_1}$ es:

$$E(\widetilde{\beta_1}) = \beta_1 + \beta_2 \frac{\sum_{i=1}^{n}(x_{i1} - \bar{x}_1)\, x_{i2}}{\sum_{i=1}^{n}(x_{i1} - \bar{x}_1)^2}$$

Donde el cociente que multiplica a β_2 es la pendiente de una regresión de x_2 sobre x_1 usando la misma muestra. El sesgo por omisión de la variable es igual a:

$$SESGO = E(\widetilde{\beta_1}) - \beta_1 = \beta_2 \frac{\sum_{i=1}^{n}(x_{i1} - \bar{x}_1)\, x_{i2}}{\sum_{i=1}^{n}(x_{i1} - \bar{x}_1)^2}$$

ii. Definición de Variables de la ENOE

Las variables que se utilizaron para formar la llave utilizada para unir las bases de datos son las siguientes:

Código	Variable
cd_a	Ciudad Representativa
ent	Entidad
con	Número de Control
n_pro_viv	Número progresivo de vivienda
v_sel	Vivienda Seleccionada
n_hog	Número de hogar
h_mud	Hogar mudado
n_ren	Número de renglón
nac_anio	Año de nacimiento
nac_mes	Mes de nacimiento
nac_dia	Día de nacimiento