

# Diferenciales salariales entre los trabajadores del sector público y del sector privado en Colombia en 1997

Darío Maldonado Carrizosa

## Abstract

*This paper uses the Roy model to measure the size of the wage differential between the workers of the public and private sector in Colombia in 1997. The use of this model is important since it eliminates the self-selection bias of the coefficients of the Mincer equations that appears when these equations are estimated separately for both sectors. It is also useful to have a reliable measure of the wage a worker will have if he would have worked in a different sector. The wage differential is measured with the traditional Oaxaca decomposition and with an additional measure of the average wage estimated differential. The paper concludes that, except for the workers with some level of graduate studies the wages in the public sector are higher than those in the private sector.*

## Resumen

El presente artículo aplica el modelo de Roy para medir el tamaño del diferencial salarial entre los trabajadores del sector público y del sector privado en Colombia en 1997. La utilización de este modelo es importante porque permite eliminar el sesgo de selección de los coeficientes de las ecuaciones de Mincer que surgen cuando se estiman ecuaciones para los dos sectores separadamente. El modelo también permite hacer predicciones apropiadas del salario contrafactual que un trabajador ganaría

---

\* Documento presentado como tesis para optar por el título de Magister en la Facultad de Economía de la Universidad de los Andes, Bogotá, 1999. El autor agradece la colaboración de su asesor Jaime Tenjo Galarza.

en el otro sector. El diferencial salarial se mide con la descomposición de Oaxaca, metodología estándar en este tipo de estudios, y con una medida adicional del diferencial promedio estimado. El estudio concluye que salvo para los trabajadores con estudios de posgrado, en el sector privado los salarios son inferiores a los del sector público.

## 1. Introducción

El diferencial salarial entre los trabajadores de los sectores público y privado es un tema de investigación que ha sido importante en muchos países (v.g. Smith, 1977, para Estados Unidos; Dustman y van Soest, 1998, para Alemania, y Asaad, 1997, para Egipto). En Colombia este es un tema que se ha estudiado poco. En la revisión bibliográfica realizada para este documento se encontraron sólo dos estudios: el primero es Psacharopoulos (1983) que no sólo estudia el tema para Colombia sino para cinco países más, y el segundo, más reciente, es Perfetti (1999).

El objeto del presente estudio es medir el diferencial salarial entre los trabajadores del sector público y los trabajadores del sector privado en Colombia. Aunque el estudio de Perfetti (1999) es reciente y bastante completo, tiene un problema especial, debido a la base de datos que utiliza, y esto es lo que motiva este trabajo.

En la literatura económica se reconoce que la decisión entre trabajar en el sector público y el sector privado es endógena, los trabajadores escogen el sector en que quieren trabajar según el salario que les ofrecen y según otras características como las compensaciones no monetarias por su trabajo. Esto genera lo que se ha llamado en la literatura sesgo de selección (Heckman, 1976 y 1979) que hace que los coeficientes obtenidos en las estimaciones de ecuaciones de salarios para los dos sectores separadamente sean sesgados. El problema del estudio de Perfetti (1999), que él mismo reconoce, se debe a que las Encuestas de Hogares no permiten corregir este problema.

En este estudio se medirá el diferencial salarial entre los trabajadores de los sectores público y privado corrigiendo el problema del sesgo de selección. Se usan los datos de la Encuesta Nacional de Calidad de Vida realizada en Colombia en 1997 a nivel nacional. La cantidad de información de esta base de datos posibilita hacer la corrección del problema mencionado.

El modelo econométrico que se usa en este estudio para medir el diferencial salarial entre trabajadores de los sectores público y privado es el que se conoce como *Switching regression model* o modelo de Roy. El modelo permite corregir el problema de selectividad en la estimación de las ecuaciones para cada uno de los sectores al endogenizar la escogencia que hace el trabajador del sector en que quiere trabajar.

Las conclusiones de este documento son similares a las de Perfetti (1999), quien encuentra evidencia para afirmar que en Colombia hay un diferencial salarial a favor de los trabajadores hombres del sector público y que el diferencial ha crecido de manera importante en los últimos años.

En este estudio se encuentra que las tasas de retorno al capital humano son diferentes en los dos sectores, y que existe un diferencial significativo, que para todos los niveles educativos favorece a los trabajadores del sector público excepto para los trabajadores con estudios de posgrado, caso en el que el diferencial favorece a los trabajadores del sector privado. Se encuentra evidencia de que el modelo que se escogió para la estimación de las ecuaciones de ingreso es el apropiado, respecto del modelo en el que no se hace la corrección de sesgo de selección, pues se rechaza la hipótesis nula de que no existe sesgo de selección. Estimar ecuaciones de ingresos como las que se hacen en este trabajo sin tener en cuenta el sesgo de selección resultaría en coeficientes sesgados.

El plan del documento es el siguiente: la primera parte es esta introducción, la segunda expone el modelo de Roy, la tercera contiene las características de la base de datos que se usará en la estimación del modelo, en la cuarta se exponen unas estadísticas descriptivas que pueden ser útiles para entender las diferencias salariales en los dos sectores, la quinta expone los resultados de la estimación de las ecuaciones de ingresos laborales, la sexta muestra la magnitud y dirección del diferencial salarial y la séptima es la conclusión del documento.

## 2. Marco teórico

El modelo que se utilizará para medir los diferenciales salariales entre el sector público y el sector privado es el modelo al que, en la literatura económica, se ha llamado *endogenous switching regression model* o mo-

delo de Roy y que es el modelo que se expone a continuación. Esta sección se fundamenta en los trabajos de Heckman (1976 y 1979), Heckman y Honoré (1990) y Maddala (1983).

El modelo que se quiere estimar se compone de dos ecuaciones equivalentes: cada una de ellas muestra cómo se determinan los salarios en cada uno de los dos sectores. Las ecuaciones pueden diferir en el valor de sus parámetros:

$$y_i = \beta_0' X_i^1 + \varepsilon_{0i} \quad \text{si } Z_i = 0 \quad (1)$$

$$y_i = \beta_1' X_i^1 + \varepsilon_{1i} \quad \text{si } Z_i = 1 \quad (2)$$

en donde  $y_i$  es el logaritmo del salario por hora del individuo  $i$ ,  $X_i^1$  es un vector con las características del capital humano del individuo  $i$  y con otras variables que afectan su salario o su productividad marginal,  $Z_i$  es una variable que indica si la persona trabaja en el sector público ( $Z_i = 0$ ) o en el sector privado ( $Z_i = 1$ ),  $\beta_0$  y  $\beta_1$  son vectores con los retornos al capital humano en cada uno de los dos sectores,  $\varepsilon_{0i}$  y  $\varepsilon_{1i}$  son los errores de cada una de las dos ecuaciones: tienen media cero y varianzas finitas y sobre éstos se tienen los siguientes supuestos:

$$E[\varepsilon_{ji}] = 0$$

$$E[\varepsilon_{ji} \varepsilon_{ji}] = \begin{cases} \sigma^2 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases} \quad (3)$$

$$\text{Cov}(X_i^1, \varepsilon_{ji}) = 0$$

para ( $i=1, 2, \dots, N$  y  $j=0, 1$ )

Se verá más adelante que debido a la base de datos que se tiene, las características del error que se observa son diferentes de las que se muestran en (3).

Una base de datos ideal para un estudio como éste debería incluir, además de las características de los trabajadores y sus empleos, el salario que cada persona gana actualmente y el que podría ganar en el otro sector. En un caso así se cumplirían los supuestos tradicionales de los errores y no sería necesario usar la técnica que se expone en esta sección del do-

umento. Como no es posible tener una base de datos con estas características, los errores de las ecuaciones de ingresos no cumplen los supuestos tradicionales.

Lo anterior hace necesario introducir correcciones a la estimación que aseguren que los errores de las ecuaciones tengan las propiedades tradicionales. El objeto de esta sección del documento es explicar la corrección que es necesario hacer a la estimación.

Los agentes no son homogéneos y escogen el sector en que quieren trabajar no aleatoriamente, por eso se dice que el modelo es endógeno; los agentes no están bajo el régimen privado o público exógenamente. La heterogeneidad de los trabajadores aparece en los errores de (1) y (2). En el modelo original sobre este tipo de problemas planteado por Roy en 1951 se dice que este es un caso en el que los agentes aprovechan las ventajas comparativas al escoger el sector en que quieren trabajar; estas ventajas comparativas surgen de la heterogeneidad de los trabajadores.

Los agentes determinan en qué sector trabajar según el diferencial salarial entre los dos sectores y un costo asociado a la decisión. Se puede suponer que existe una función de costo para los agentes como la siguiente:

$$c_i = \alpha' X_i^2 + u_i \quad (4)$$

en donde  $c_i$  es el costo para la persona  $i$  de tomar la decisión,  $X_i^2$  es un vector con las características de la persona  $i$  que afectan el costo de tomar la decisión y que pueden ser diferentes de las que se incluyen en el vector  $X_i^1$ ,  $\alpha$  es un vector con los coeficientes que relacionan las características y el costo y  $u_i$  es el error que tiene función de distribución normal con media cero. Con

$$Z_i^* = y_i(Z_i = 0) - y_i(Z_i = 1) - c_i \quad (5)$$

Una persona decide trabajar en el sector público si  $Z^*$  es positiva y decide trabajar en el sector privado si  $Z^*$  es negativa. Usando (1), (2) y (5) se obtiene la siguiente expresión:

$$Z_i^* = (\beta_0 - \beta_1)' X_i^1 - \alpha' X_i^2 + (\varepsilon_{0i} - \varepsilon_{1i} - u_i) \Rightarrow Z_i = \begin{cases} 0 & \text{si } Z_i^* > 0 \Leftrightarrow \eta_i > \pi' X_i \\ 1 & \text{si } Z_i^* < 0 \Leftrightarrow \eta_i < \pi' X_i \end{cases} \quad (6)$$

en donde  $\pi = [(\beta_0 - \beta_1) - \alpha]$ ,  $\eta_i = \varepsilon_{0i} - \varepsilon_{1i} + u_i$  y  $X_i = [X_i^1 | X_i^2]$

El modelo completo se compone de tres ecuaciones: (1), (2) y (6). En este contexto surgen varios problemas: el primero es responder a la pregunta de por qué no se pueden estimar (1) y (2) separada e independiente de (6), y el segundo es cómo estimar (6) cuando no se conoce  $Z_i^*$  sino  $Z_i$ .

La respuesta al primer problema se encuentra en Heckman (1976 y 1979) en donde se muestra que si se estima un ecuación como (1) o como (2), en donde los agentes han escogido no aleatoriamente si pertenecen o no a la muestra (si trabajan o no en determinado sector), los coeficientes que se encuentran por mínimos cuadrados ordinarios son sesgados. Heckman (1979) caracteriza el problema como uno de variables omitidas.

Si se opera (1) con el operador de expectativas, teniendo en cuenta que en la base de datos que se tiene sólo se conoce el salario de la persona en el sector en que trabaja, y el salario en el otro sector no es observable, se obtiene:

$$E[y_i | Z_i = 0] = \beta_0' X_i + E[\varepsilon_{0i} | Z_i = 0] \quad (7)$$

para que la estimación de (1) o de (2) pueda hacerse por MCO  $E[\varepsilon_{0i} | Z_i = 0]$  debería ser igual a cero; el problema es que como  $\varepsilon_{0i}$  y  $\eta_i$  están correlacionados, éste no será cero. Usando (6) y (7):

$$\begin{aligned} E[\varepsilon_{0i} | Z_i = 0] &= E[\varepsilon_{0i} | \pi' X_i \geq \eta_i] = E[\sigma_{0\eta} \eta_i | \pi' X_i \geq \eta_i] \\ &= E[\sigma_{0\eta} \eta_i | \pi_i' X_i \geq \eta_i] = -\frac{\sigma_{0\eta}}{\sigma_\eta} \lambda_{0i} \quad (8) \\ \text{con } \lambda_{0i} &= \frac{\phi(\Theta_i)}{\Phi(\Theta_i)} \quad \text{y} \quad \Theta_i = \frac{\pi_i' X_i}{\sigma_\eta} \end{aligned}$$

en donde se ha tenido en cuenta que, por las propiedades de la función de distribución normal bivariada, y dado que  $Z_i = 0$ ,  $\varepsilon_{0i}$  tiene media  $\sigma_{0\eta} \eta_i$  y propiedades de las distribuciones normales truncadas. En (8),  $\phi(\cdot)$  y  $\Phi(\cdot)$  son la función de densidad y la función de probabilidad de la distribución normal estándar, respectivamente.  $\sigma_\eta$  es la desviación estándar de  $\eta_i$ , y  $\sigma_{0\eta}$  es la covarianza de  $\varepsilon_{0i}$  y  $\eta_i$ .

Si se hace el mismo proceso para la ecuación (2) se obtiene:

$$E[y_i|Z_i = 1] = \beta_1' X_i + E[\varepsilon_{1i}|Z_i = 1]$$

$$E[\varepsilon_{1i}|Z_i = 1] = E[\varepsilon_{1i}|\pi' X_i < \eta_i] = \frac{\sigma_{1\eta}}{\sigma_\eta} \lambda_{1i} \quad (9)$$

$$\text{con } \lambda_{1i} = \frac{\phi(\Theta_i)}{1 - \Phi(\Theta_i)} \quad \text{y} \quad \Theta_i = \frac{\pi' X_i}{\sigma_\eta}$$

en donde  $\sigma_{1\eta}$  es la covarianza de  $\varepsilon_{1i}$  y  $\eta_i$ .

En (8) y (9) se observa que las medias de los errores de (1) y (2), dado que ya se ha escogido un sector en el cual trabajar, son cero solamente si  $\sigma_{0\eta}$  y  $\sigma_{1\eta}$  son iguales a cero, lo que no ocurre en este caso dada la estructura supuesta de los errores. Si se estiman por MCO las ecuaciones (1) y (2) se obtienen coeficientes sesgados; este problema es el conocido en la literatura como sesgo muestral o sesgo de selección.

Usando los resultados de (8) en (7) y (9), se obtienen dos expresiones que sí son estimables mediante MCO sin el problema de sesgo muestral:

$$y_i = \beta_0' X_i - \frac{\sigma_{0\eta}}{\sigma_\eta} \lambda_{0i} + \varepsilon_{0i} \quad \text{si } Z_i = 0 \quad (10)$$

$$y_i = \beta_1' X_i + \frac{\sigma_{1\eta}}{\sigma_\eta} \lambda_{1i} + \varepsilon_{1i} \quad \text{si } Z_i = 1 \quad (11)$$

donde

$$\varepsilon_{0i} = \varepsilon_{1i} + \frac{\sigma_{1\eta}}{\sigma_\eta} \lambda_{1i} \quad (12)$$

$$\varepsilon_{1i} = \varepsilon_{1i} - \frac{\sigma_{1\eta}}{\sigma_\eta} \lambda_{1i} \quad (13)$$

Como se dijo, en Heckman (1976 y 1979) se caracteriza este problema como uno de variables omitidas. El sesgo muestral al estimar (1) y (2) separada e independientemente sin tener en cuenta el proceso de selección entre los dos regímenes surge de omitir la variable  $\lambda_{0i}$  en la ecuación (1) o la variable  $\lambda_{1i}$  en la ecuación (2).

En este modelo los agentes escogen el sector en que quieren trabajar de manera que puedan aprovechar sus ventajas comparativas. La manera

como éstas se aprovechan depende de los signos de  $\sigma_{0\eta}$  y  $\sigma_{1\eta}$ . Si  $\sigma_{0\eta} < 0$  y  $\sigma_{1\eta} > 0$ , la media condicional del ingreso de los trabajadores en cada uno de los dos sectores sería superior a la media incondicional de su sector; por sus ventajas comparativas el promedio del salario contrafactual (i.e., el que ganarían en el otro sector) es menor a la media incondicional (para  $Z_i = 0$  es  $\beta_0' X_i^1$  y para  $Z_i = 1$  es  $\beta_1' X_i^1$ ) para los dos grupos de trabajadores. Si  $\sigma_{0\eta} < 0$  y  $\sigma_{1\eta} < 0$ , la media condicional del ingreso de los trabajadores del sector 0 sería superior pero el de los trabajadores del sector 1 sería inferior al salario medio condicional de cada sector. En este caso la media del salario contrafactual condicional de los trabajadores del sector 0 es mayor a la media incondicional del sector 1 y el salario medio contrafactual de los trabajadores del sector 1 sería menor al salario medio incondicional en el sector 0. Si  $\sigma_{0\eta} > 0$  y  $\sigma_{1\eta} > 0$  el caso es el inverso al anterior. El caso en que  $\sigma_{0\eta} > 0$  y  $\sigma_{1\eta} < 0$  es imposible dada la definición de  $\eta_i$  y las consecuencias de ésta sobre las covarianzas.

Desde el punto de vista económico lo que se vería en un mercado como este es que por lo menos uno de los dos grupos de trabajadores recibiría salarios superiores al salario promedio del mercado. Esto se ve claramente al detallar el último caso expuesto, el cual es imposible desde el punto de vista económico porque significaría que los trabajadores no aprovechan sus ventajas comparativas, en la medida en que el salario promedio de cada uno de los dos grupos podría ser mayor si trabajaran en el otro sector. Con el supuesto de trabajadores racionales este es un caso que no se observaría nunca.

La razón por la que un grupo de trabajadores puede tener un salario medio condicional que sea inferior al salario medio incondicional es que, por el aprovechamiento de las ventajas comparativas, los trabajadores que pueden recibir salarios más altos en los dos sectores se concentran en uno solo de ellos y en el otro sector se concentran los trabajadores que podrían recibir salarios menores en los dos sectores.

Un problema mencionado antes era el de cómo estimar la ecuación (6) cuando no se conoce  $Z_i^*$ . Para esto se usa la variable cualitativa  $Z_i$ , que sí es conocida, y se obtiene el vector de parámetros  $\pi$  de la estimación de un modelo probit. Con las estimaciones de estos parámetros se pueden obtener valores estimados para  $\phi(\cdot)$  y  $\Phi(\cdot)$ , y usar estos valores estimados para estimar (10) y (11) mediante MCO.



La implementación empírica de las dos formas de estimar el modelo desarrollado requiere de varias precisiones adicionales; una es cómo manejar el problema de la heterocedasticidad de los errores, y la otra es cuáles variables se deben incluir en las dos ecuaciones.

Las ecuaciones siguientes muestran que la matriz de varianzas y covarianzas de este modelo no cumple los supuestos tradicionales:

$$\begin{aligned}
 E(\epsilon_{0i}^2) &= \sigma_0^2 [(1 - \rho_0^2) + \rho_0^2 (1 + \Theta_i \lambda_{0i} - \lambda_{0i}^2)] \\
 E(\epsilon_{1i}^2) &= \sigma_1^2 [(1 - \rho_1^2) + \rho_1^2 (1 + \Theta_i \lambda_{1i} - \lambda_{1i}^2)] \\
 \text{con } \rho_0 &= \frac{\sigma_{0\eta}}{\sigma_0 \sigma_\eta} \quad \text{y} \quad \rho_1 = \frac{\sigma_{1\eta}}{\sigma_1 \sigma_\eta}
 \end{aligned} \tag{14}$$

Estas expresiones se obtienen de las propiedades de la función de distribución normal bivariada. Lo que muestran las expresiones en (14) es que la varianza de los errores de las ecuaciones de ingresos no es constante, por lo que la estimación del modelo de Roy debe hacerse por mínimos cuadrados ponderados según la estructura señalada por ellas.

Para corregir el problema de heterocedasticidad de las ecuaciones de ingresos con la corrección del sesgo de selección es necesario tener valores estimados de los parámetros  $\sigma_i^2$  y  $\rho_i$  con los que se puedan construir las ponderaciones apropiadas. Para ello se puede usar la estimación del parámetro asociado al  $\lambda_i$  en una estimación de la ecuación de ingresos por mínimos cuadrados ordinarios y tener en cuenta que este parámetro es igual a  $\rho_i \sigma_i^2$ . Usando el valor estimado que se obtiene en la misma estimación de  $\sigma_i^2$  se obtiene el parámetro  $\rho_i$ . Con estos valores se pueden construir las ponderaciones para hacer la estimación por medio de mínimos cuadrados ponderados.

Una manera alternativa para obtener estos parámetros es la que se plantea en Heckman (1976). En este estudio se sugiere usar una regresión auxiliar en la que la variable dependiente es el error de la ecuación del sector 0 estimada por MCO y las variables dependientes son  $\Theta_i \lambda_i - \lambda_i^2$  y una constante; el parámetro asociado a la constante será una estimación de  $\sigma_0^2$ , el otro parámetro será una estimación de  $\sigma_0^2 \rho_0^2$ .

En este documento se utiliza la primera de las alternativas planteadas. Con ellas hay un problema adicional y es que en ninguna existe la forma de asegurar que la estimación del parámetro  $\rho_i$  se encuentre en el intervalo unitario, ni que asegure que los valores de las varianzas estimadas sean positivos. Además en la alternativa propuesta por Heckman (1976) es probable que la estimación de  $\sigma_i^2$  sea negativa.

En el modelo que se acaba de exponer se supuso que los agentes escogen únicamente entre dos alternativas. Este es un supuesto que, en el marco del objetivo de este trabajo, no se cumple. Los agentes en el mercado laboral tienen un rango mayor de opciones que comienzan con la decisión de participación y pueden ir hasta decisiones mucho más específicas que la decisión sector público-sector privado. Esta consideración cambia las consecuencias del modelo respecto de la estimación y respecto de la interpretación de las estimaciones de las covarianzas. Respecto del primero de estos puntos lo que hay que decir es que si se mantiene el supuesto de la normalidad de los errores de las ecuaciones de selección y de ingresos, la manera de estimar el modelo no cambia, lo único que cambia es el grupo de personas que se usa para estimar los  $\lambda$ . Sobre el segundo punto es importante tener en cuenta que al existir más de dos sectores es posible que los trabajadores de los dos sectores en los que estamos interesados obtengan salarios promedio inferiores a los salarios promedio incondicionales de su sector. Cuando se explicaron las consecuencias de los posibles signos de las covarianzas entre el error de la ecuación de selección y el de la ecuación de ingresos, se dijo que no era factible el caso en el que los dos grupos de trabajadores tenían salarios promedio inferiores a los salarios promedio incondicionales de su sector. Sin embargo, cuando se piensa que hay más de dos posibilidades de selección, ese caso es factible. Esto se debe a la alternativa de que un grupo de trabajadores pueda ser el que obtiene los salarios más altos en cualquiera de los sectores pero que se encuentre concentrado en uno solo de ellos.

### 3. Los datos

Como se mencionó en la introducción, la fuente más importante de información para este documento es la Encuesta Nacional de Calidad de Vida (ENCV) realizada por el Departamento Administrativo Nacional de Estadística (DANE) en 1997. La encuesta es de corte transversal y se realizó a nivel nacional. Se utiliza en este estudio porque contiene informa-

ción importante de las personas que no se encuentra en la Encuesta Nacional de Hogares, que es la que se usa tradicionalmente para estudios de este tipo.

La ENCV se aplicó a nivel nacional a 9.121 hogares equivalentes a 38.512 personas. Con los factores de expansión esto equivale a 9.436.091 hogares y 39.842.424 personas. En la encuesta hay 14.799 trabajadores de los cuales 4.095 son empleados asalariados del sector privado y 923 son empleados del sector público. Con factores de expansión esto equivale a 15.559.270 trabajadores en total, 5.120.976 asalariados del sector privado y 1.003.925 trabajadores del sector público.

Según el documento en el que se comenta el diseño de la muestra de la ENCV (Gómez, 1998) en una muestra de 8.000 hogares con un esrel de 10% y un defl de 1.5 es factible analizar fenómenos con una ocurrencia de 5% en 12 desagregaciones, fenómenos con una ocurrencia de 10% en 26 desagregaciones y fenómenos con una ocurrencia de 20% en 60 desagregaciones. Esto dice que al ser los trabajadores del sector público el 6.45% del total de los trabajadores es posible trabajar en 12 desagregaciones. En ninguna parte de este trabajo se hacen más de ocho desagregaciones de este grupo de trabajadores, lo que permite decir que la ENCV es apropiada para el estudio que se ha hecho.

Además de la ENCV se usaron algunas etapas de la Encuesta Nacional de Hogares (ENH) para construir estadísticas descriptivas de los dos sectores en cuestión. Esto se hizo debido a que la ENH tiene información sobre las características de los empleos que no tiene la ENCV.

La ENH de 1998 (etapa 100) se hizo a 21.478 hogares en los que hay 89.712 personas. Expandiendo, esto equivale a 3.726.368 hogares y 15.564.035 personas. En la encuesta hay 35.534 trabajadores de los cuales 16.609 son asalariados del sector privado y 3.198 son trabajadores del sector público. Expandiendo, esto equivale a 6.322.171 trabajadores en total, 3.207.920 trabajadores asalariados del sector privado y 548.630 trabajadores del sector público.

La ENH de 1996 (etapa 92) consta de 19.520 hogares u 82.806 personas. De éstas 32.452 son trabajadores, 17.206 son trabajadores asalariados del sector privado y 2.837 son trabajadores del sector público. Expandiendo,

esto equivale a 3.131.747 hogares, 13.285.217 personas, 5.353.438 trabajadores, 2.957.786 trabajadores asalariados del sector privado y 449.899 trabajadores del sector público.

Como se ve el tamaño de la ENH es mayor que el de la ENCV por lo que las dudas sobre su utilidad para este estudio son menores.

#### **4. Algunas características de los empleados y los empleos del sector público y del sector privado**

En el Cuadro 1 aparecen algunas estadísticas descriptivas de los trabajadores y los empleos de los sectores público y privado. Las estadísticas, que se elaboraron teniendo en cuenta los factores de expansión de la encuesta, permiten ver que tanto los trabajadores como los empleos en los dos sectores son diferentes. Además permiten pensar en posibles hipótesis que expliquen las diferencias salariales entre los dos sectores. Aunque no es el objetivo principal de este documento, vale la pena mencionar algunos aspectos de ellas.

Los trabajadores del sector público tienen un nivel educativo más alto y mayor promedio de edad que los del sector privado. La primera hipótesis podría ser que las posibles diferencias en los salarios se deben a características en cuanto a la acumulación de capital humano en los dos sectores y favorecería a los trabajadores del sector público. A favor de encontrar salarios promedio superiores en el sector privado se ve en el cuadro que la proporción de mujeres en el sector privado es menor, lo que, teniendo en cuenta que en Colombia existe discriminación salarial en contra de las mujeres (Tenjo, 1992), puede llevar a que se encuentre que el salario promedio en el sector privado es superior al del sector público. Lo que esto dice es que la medición del diferencial salarial debe hacerse de manera que se aisle tanto el efecto del género como el del capital humano de dicho diferencial.

Una de las hipótesis importantes para explicar diferenciales salariales entre trabajadores homogéneos es la llamada hipótesis de diferenciales compensatorios. Estrictamente, la medición del diferencial salarial debería ha-

cerse teniendo en cuenta las características del empleo, de manera que el diferencial no esconda condiciones desfavorables del empleo en cualquiera de los dos sectores. Este tipo de control no se tendrá en este documento pero vale la pena mencionar que, de acuerdo con el Cuadro 1, los trabajadores del sector público tienen mejores condiciones laborales que los del sector privado, las proporciones de trabajadores con afiliación a seguridad social en salud y pensión y de trabajadores con contratos a término indefinido son mayores en el sector público que en el privado. Lo que la hipótesis de diferenciales compensatorios predice frente a este hecho es que probablemente los empleados del sector privado tienen mayores salarios que los compensan por las condiciones desfavorables de sus empleos: la predicción en un modelo con competencia imperfecta, v.g. con restricciones a la entrada al sector público, podría ser diferente.

En el Cuadro 2, en el que aparecen los promedios del salario por hora para los distintos niveles educativos, se observa que hay evidencia para afirmar que existe un diferencial salarial a favor de los trabajadores del sector público.

En los Cuadros 1 y 2 se resumen las mismas características expuestas pero usando la muestra de empleados de empresas con más de once trabajadores. Ninguna de las conclusiones anteriores cambia con respecto a las que se obtuvieron con las estadísticas con toda la muestra de empleados, aunque la magnitud de las diferencias es menor. La razón por la que se elaboraron estas estadísticas es que uno de los posibles determinantes del salario es el tamaño de la planta: se reconoce una relación directa entre el salario y el tamaño de planta.

Como parte de este ejercicio se hicieron estadísticas similares usando dos etapas de la ENH correspondientes a junio de los años 1996 y 1998, que fueron escogidas porque tienen más información que las encuestas de otros meses y por su cercanía al año en que se realizó la ENCV usada en este estudio. La ventaja de usar estas encuestas es que en ellas hay información que no contiene la ENCV sobre todo respecto de las características del empleo. Además, la ENH es la fuente de información usada tradicionalmente en los estudios sobre el mercado laboral colombiano y vale la pena comparar algunos resultados obtenidos con la ENCV con los de la ENH. Los resultados mencionados aparecen en los Cuadros 3 a 6. Se reportan las mismas estadísticas que se reportaron usando la ENCV, además de estadísticas relativas al subempleo.

**Cuadro 1 Estadísticas descriptivas de las características de los trabajadores de los sectores público y privado según la ENCV, 1997**

	Toda la muestra		Muestra con empleados de empresas de más de once empleados	
	Sector privado	Sector público	Sector privado	Sector público
Proporción dentro de los trabajadores asalariados	81,01%	18,99%	71,69%	28,31%
<b>Características de los trabajadores</b>				
Promedio de edad	33,12	37,51	33,9986	37,5360
Promedio del número de años de educación	8,60	11,67	9,4193	11,6737
Proporción de mujeres	38,48%	52,94%	38,92%	52,95%
Proporción de personas sin educación	2,97%	0,40%	1,72%	0,40%
Proporción de personas con primaria	27,15%	11,86%	22,52%	11,76%
Proporción de personas con secundaria	52,51%	45,98%	52,28%	46,10%
Proporción de personas con educación técnica	5,58%	7,11%	6,80%	7,00%
Proporción de personas con universidad sin título	3,37%	4,35%	4,50%	4,36%
Proporción de personas con universidad con título	7,01%	21,74%	10,11%	21,80%
Proporción de personas con posgrado	1,29%	8,56%	2,11%	8,59%
<b>Características del empleo</b>				
Proporción de personas con seguridad social en salud	78,06%	95,75%	89,96%	95,96%
Proporción de personas con seguridad social en pensión	51,12%	84,31%	69,35%	84,50%
Proporción de trabajadores con contrato a término indefinido	50,06%	87,35%	77,03%	87,45%

Entre paréntesis se presentan los valores absolutos de los estadísticos t para la prueba de hipótesis de igualdad de las estadísticas correspondientes para los dos sectores bajo el supuesto de varianzas diferentes.

\* significativo al 5%.

Fuente: cálculos del autor con datos de la Encuesta Nacional de Calidad de Vida, DANE, Colombia, 1997.

**Cuadro 2** Promedio de los ingresos laborales por hora según la ENCV, 1997

Nivel educativo	Toda la muestra			Muestra con empleados de empresas de más de once empleados		
	Sector privado	Sector público	Ttest	Sector privado	Sector público	Ttest
Ningún año	889.77	1223.54	(1.22)	1187.16	1223.54	(0.12)
Primaria	1226.77	2061.13	(3.02) *	1443.47	2079.23	(2.26) *
Secundaria	1738.61	2725.59	(6.66) *	2036.89	2725.59	(4.24) *
Técnica	2962.25	3006.40	(0.11)	3272.87	2995.84	(0.75)
Universitaria sin título	3931.41	4551.14	(0.72)	4208.50	4554.14	(0.38)
Universitaria con título	5655.50	6027.52	(0.75)	5902.75	6027.52	(0.24)
Posgrado	9679.64	6182.01	(3.19) *	9350.83	6182.01	(2.92) *

Entre paréntesis se presentan los valores absolutos de los estadísticos t para la prueba de hipótesis de igualdad de las estadísticas correspondientes para los dos sectores bajo el supuesto de varianzas diferentes.

\* significativo al 5%.

Fuente: cálculos del autor con datos de la Encuesta Nacional de Calidad de Vida, DANE, Colombia, 1997.

Respecto de las características de las que se habló con base en la ENCV, la conclusión no cambia: los trabajadores del sector público tienen mayor edad, un nivel educativo más alto y mejores condiciones laborales que los trabajadores del sector privado.

En la ENH se define subempleado a un trabajador que afirma querer trabajar más tiempo o querer un trabajo diferente al que tiene actualmente. El subempleo es entonces un indicador de las condiciones laborales a las que los trabajadores están sujetos, por lo tanto si se encuentra que un determinado sector tiene un subempleo alto comparado con otro se podría pensar que en ese primer sector las condiciones laborales son peores que en el segundo. Lo que se encuentra es que la proporción de trabajadores que dice estar subempleado es menor en el sector público, para las dos encuestas consideradas. Más aún, al considerar las razones del subempleo se encuentra que en el sector público una menor proporción de los subempleados dice serlo por bajos ingresos o porque no les gusta el empleo.

**Cuadro 3** Estadísticas descriptivas de las características de los trabajadores de los sectores público y privado según la ENH de junio de 1996

	Toda la muestra		Muestra con empleados de empresas de más de once empleados			
	Sector privado	Sector público	Ttest	Sector privado	Sector público	Ttest
Proporción dentro de los trabajadores asalariados	86.79%	13.21%		81.08%	18.92%	
<b>Características de los trabajadores</b>						
Promedio de edad	32.33	36.77	(20.95)*	32.77	36.77	(18.01)*
Promedio del número de años de educación	9.28	12.64	(12.34)*	9.97	12.65	(31.11)*
Proporción de mujeres	39.50%	46.65%	(8.94)*	40.43%	46.61%	(6.68)*
Proporción de personas sin educación	1.14%	0.16%	(5.43)*	0.76%	0.16%	(3.97)*
Proporción de personas con primaria incompleta	7.28%	1.70%	(11.74)*	5.69%	1.62%	(8.53)*
Proporción de personas con primaria completa	15.08%	5.11%	(11.74)*	12.61%	5.10%	(11.13)*
Proporción de personas con secundaria incompleta	27.87%	10.46%	(19.31)*	24.76%	10.19%	(16.14)*
Proporción de personas con secundaria completa	26.80%	29.22%	(4.88)*	28.61%	29.44%	(1.93)*
Proporción de personas con menos de cinco años de universidad	10.35%	15.63%	(11.90)*	12.04%	15.60%	(7.40)*
Proporción de personas con más de cinco años de universidad	10.98%	37.34%	(36.78)*	15.00%	37.50%	(25.88)*
<b>Características del empleo</b>						
Proporción de personas con seguridad social en salud	67.22%	94.25%	(32.72)*	82.04%	94.50%	(17.11)*
Proporción de personas con seguridad social en pensión	46.27%	81.73%	(35.89)*	61.49%	82.36%	(20.69)*
Proporción de trabajadores temporales	81.74%	90.12%	(12.143)*	85.61%	90.49%	(7.15)*
Proporción de subempleados	20.41%	11.13%	(11.87)*	17.64%	11.06%	(7.55)*
Proporción de subempleados por bajos ingresos	60.73%	46.45%	(10.82)*	62.11%	45.72%	(7.81)*
Proporción de subempleados porque no les gusta el trabajo	6.66%	3.40%	(4.09)*	5.60%	3.53%	(2.64)*

Entre paréntesis se presentan los valores absolutos de los estadísticos t para la prueba de hipótesis de igualdad de las estadísticas correspondientes para los dos sectores bajo el supuesto de varianzas diferentes.

\* significativo al 5%.  
Fuente: cálculos del autor con datos de la Encuesta Nacional de Hogares, etapa 92, DANE, Colombia, junio 1996.



**Cuadro 4 Estadísticas descriptivas de las características de los trabajadores de los sectores público y privado según la ENH de junio de 1998**

	Toda la muestra		Muestra con empleados de empresas de más de once empleados	
	Sector privado	Sector público	Sector privado	Sector público
Proporción dentro de los trabajadores asalariados	85,38%	14,62%	79,27%	20,73%
<b>Características de los trabajadores</b>				
Promedio de edad	32,78	37,60	33,32	37,68
Promedio del número de años de educación	9,88	13,03	10,70	13,12
Proporción de mujeres	41,66%	50,52%	42,28%	50,15%
Proporción de personas sin educación	0,95%	0,33%	0,50%	0,23%
Proporción de personas con primaria incompleta	6,57%	1,21%	4,17%	1,23%
Proporción de personas con primaria completa	13,14%	4,42%	11,42%	1,32%
Proporción de personas con secundaria incompleta	23,48%	7,87%	19,75%	7,32%
Proporción de personas con secundaria completa	29,85%	30,19%	31,22%	30,05%
Proporción de personas con menos de cinco años de universidad	11,25%	17,35%	13,75%	17,25%
Proporción de personas con más de cinco años de universidad	14,16%	38,48%	18,91%	39,43%
<b>Características del empleo</b>				
Proporción de personas con seguridad social en salud	68,63%	94,13%	83,59%	87,22%
Proporción de personas con seguridad social en pensión	49,03%	86,68%	67,53%	87,15%
Proporción de trabajadores temporales	76,82%	87,73%	81,65%	88,26%
Proporción de subempleados	26,56%	13,74%	22,71%	13,51%
Proporción de subempleados por bajos ingresos	64,93%	58,05%	64,88%	59,24%
Proporción de subempleados porque no les gusta el trabajo	4,20%	2,59%	3,97%	2,78%

Entre paréntesis se presentan los valores absolutos de los estadísticos t para la prueba de hipótesis de igualdad de las estadísticas correspondientes para los dos sectores bajo el supuesto de varianzas diferentes.

\* significativo al 5%.

Fuente: cálculos del autor con datos de la Encuesta Nacional de Hogares, etapa 100, DANE, Colombia, junio 1998.

**Cuadro 5 Promedio de los ingresos laborales por hora, 1996**

Nivel educativo	Toda la muestra			Muestra con empleados de empresas de más de once empleados		
	Sector privado	Sector público	Ttest	Sector privado	Sector público	Ttest
Ningún año o preescolar	924.90	1347.64	-0.7249	963.22	1347.64	(0.6695)
Primaria incompleta	895.94	1009.75	(2.4558) *	955.93	1028.55	(1.7831)
Primaria completa	1006.44	1321.13	(2.1859) *	1079.86	1337.92	(2.2723) *
Secundaria incompleta	1050.81	1512.55	(9.2821) *	1116.03	1529.80	(6.9063) *
Secundaria completa	1284.45	1716.30	(10.5296) *	1365.03	1713.73	(7.4347) *
Menos de cinco años de educación superior	1934.57	2515.39	(5.9398) *	2055.33	2514.60	(4.1653) *
Más de cinco años de educación superior	4265.73	4339.40	(0.4236)	4439.06	4371.55	(0.5329)

Entre paréntesis se presentan los valores absolutos de los estadísticos t para la prueba de hipótesis de igualdad de las estadísticas correspondientes para los dos sectores bajo el supuesto de varianzas diferentes.

\* significativo al 5%.

Fuente: cálculos del autor con datos de la Encuesta Nacional de Hogares, etapa 92, DANE, Colombia, junio 1996.

**Cuadro 6 Promedio de los ingresos laborales por hora, 1998**

Nivel educativo	Toda la muestra			Muestra con empleados de empresas de más de once empleados		
	Sector privado	Sector público	Ttest	Sector privado	Sector público	Ttest
Ningún año o preescolar	1228.15	1623.10	(3.0056) *	1283.06	1579.95	(2.7976) *
Primaria incompleta	1200.59	1493.72	(3.0733) *	1294.48	1428.42	(1.3206)
Primaria completa	1329.33	1572.55	(3.9909) *	1426.82	1579.31	(3.2944) *
Secundaria incompleta	1434.17	2312.46	(11.3359) *	1511.99	2329.36	(6.648) *
Secundaria completa	1859.97	2572.80	(14.4607) *	1984.44	2600.30	(11.7565) *
Menos de cinco años de educación superior	2955.24	3772.66	(6.6359) *	3139.91	3678.59	(4.9535) *
Más de cinco años de educación superior	7238.55	6638.77	(0.3423)	7353.08	6673.53	(0.2606) *

Entre paréntesis se presentan los valores absolutos de los estadísticos t para la prueba de hipótesis de igualdad de las estadísticas correspondientes para los dos sectores bajo el supuesto de varianzas diferentes.

\* significativo al 5%.

Fuente: cálculos del autor con datos de la Encuesta Nacional de Hogares, etapa 100, DANE, Colombia, junio 1998.

Además de las estadísticas descriptivas mencionadas se estimaron modelos probit para encontrar las diferencias entre los trabajadores asalariados del sector privado y los trabajadores del sector público. Estos modelos se reportan en el Cuadro 7, en donde no se presentan los coeficientes de la estimación, sino las derivadas de la probabilidad de trabajar en el sector público con respecto a la variable correspondiente para las variables continuas y el cambio discreto en la misma probabilidad generado por un cambio en las variables dummy.

**Cuadro 7** Modelo probit para la decisión de trabajar en el sector público o en el sector privado

	Sin separar sexos	Hombres	Mujeres
Dummy: mujer	0.3790 (6.46) *		
Dummy: lugar de nacimiento	0.1522 (2.46) *	0.1427 (1.74)	0.1589 (1.65)
Años de educación	0.1257 (15.73) *	0.1176 (13.41) *	0.1372 (12.94) *
Años de experiencia potencial	0.0335 (4.38) *	0.0273 (2.59) *	0.0334 (2.83) *
Años de experiencia potencial al cuadrado	-0.0003 (2.15) *	-0.0002 (1.31)	-0.0002 (0.87)
Dummy: casado o en unión libre	0.0984 (1.78)	-0.0500 (0.48)	0.2270 (2.69) *
Dummy: jefe de hogar	0.2310 (3.75) *	0.3216 (2.92) *	0.3508 (3.53) *
Dummy: migrante	0.0022 (0.03)	0.0317 (0.35)	-0.0389 (0.34)
Dummy: residente de centro poblado	0.6149 (7.86) *	0.4625 (4.52) *	0.8224 (6.63) *
Dummy: residente área rural dispersa	0.3199 (3.38) *	0.0841 (0.71)	0.8034 (4.80) *
Constante	-3.1897 (25.39) *	-2.9246 (18.38) *	-3.1222 (16.14) *
Observaciones	4312	2615	1697

En la tabla se presenta la derivada de la función de probabilidad con respecto a la variable para las variables continuas, y el cambio generado por un cambio discreto para las variable dummy.

Valores absolutos de los estadísticos z en paréntesis.

\* significativo al 5%.

Fuente: cálculos del autor con datos de la Encuesta Nacional de Calidad de Vida, DANE, Colombia, 1997.

En estas estimaciones la variable dependiente es una dummy que toma el valor de uno si la persona trabaja en el sector público y de cero si trabaja en el sector privado; la muestra utilizada para el ejercicio incluye solamente a los trabajadores asalariados (se excluye a los trabajadores independientes y a los patrones o empleadores).

Se incluyeron como variables explicativas de la probabilidad de trabajar en el sector público: una variable dummy para el sexo, que toma valor de uno si la persona es mujer y de cero si es hombre; el número de años de educación; el número de años de experiencia potencial en forma cuadrática; variables dummy que toman el valor de uno si la persona está casada o en unión libre, si es migrante (no nació en ese municipio y lleva viviendo en él menos de seis años), si vive en un centro poblado, si vive en área rural (respecto de estas dos últimas, la característica de control es si la persona vive en la cabecera de un municipio), si es jefe de hogar y cero de lo contrario; se incluyó también un variable dummy que toma valor de cero si la persona nació en un centro urbano con alcaldía municipal y de uno si nació en un corregimiento, vereda, inspección de policía, caserío o en el campo.

Los coeficientes del Cuadro 7 se interpretan de la siguiente manera: para las variables continuas, como el número de puntos porcentuales en que cambiaría la probabilidad de que una persona se encuentre trabajando en el sector público ante un cambio de una unidad en la variable, y para las variables dummy, como la diferencia entre la probabilidad de estar en el sector público entre una persona con la característica señalada por la variable y una persona que no tiene la característica.

Las conclusiones que se obtuvieron en este ejercicio son similares a las que se obtienen con las estadísticas descriptivas. El número de años de educación incrementa significativamente la probabilidad de trabajar en el sector público. El número de años de experiencia también afecta positivamente la probabilidad de trabajar en el sector público. Aunque se predice un comportamiento cuadrático, el máximo se encuentra para los tres casos, fuera del rango relevante (por encima de los cien años de experiencia). En estas estimaciones se ve también que la probabilidad de trabajar en el sector público es mayor para las mujeres, para las personas que nacieron en corregimientos, veredas, inspecciones de policía, caseríos o en el campo, para los residentes de centros poblados y áreas rurales y para los jefes de hogar, respecto de sus respectivas alternativas.

## 5. Resultados de la estimación

Para cumplir con el objetivo de medir el diferencial salarial entre los trabajadores del sector se estimó el modelo del que se habló en la segunda parte del trabajo: específicamente se estimaron las ecuaciones (10) y (11). La estimación se hizo con el método mínimos cuadrados en dos etapas con la corrección apropiada de la matriz de covarianzas.

Los pasos de la estimación son los siguientes: se estimó el *probit* y con él se estimó la variable  $\lambda_i$ , después se estimó la ecuación de ingresos incluyendo como regresor el  $\lambda_i$ . De esta regresión se extrajeron los parámetros  $\rho_i$  y  $\sigma_i^2$ , y con éstos se construyeron las ponderaciones que se usan para la estimación de la ecuación de ingreso por mínimos cuadrados generalizados.

Se hicieron estimaciones para hombres y mujeres por separado y para el grupo con hombres y mujeres juntos. La más confiable es la estimación para los hombres ya que en el caso de éstos la decisión de participar o no en el mercado laboral no es realmente importante. En el caso de las mujeres, ésta es una decisión importante y es posible que se introduzca un nuevo sesgo de selección en la ecuación. Además, como se explicará más adelante, la manera como se construye la variable para la experiencia es apropiada para los hombres pero no para las mujeres lo que hace menos confiable la estimación para este último grupo de trabajadores.

La escogencia de las variables que se incluyeron en la estimación responde tanto a necesidades teóricas como a las posibilidades que da la base de datos utilizada. Se estimaron dos tipos de ecuaciones que difieren en las variables independientes incluidas en ellas. La variable dependiente es el logaritmo del salario por hora. La primera ecuación es la utilizada tradicionalmente en estudios sobre retornos al capital humano, las variables independientes son los años de educación formal, los años de experiencia, los años de experiencia al cuadrado y una variable *dummy* para el sexo en la estimación con la muestra de hombres y mujeres. La segunda ecuación incluye un *spline* que permite tener en cuenta la posibilidad de que los distintos niveles o tipos de educación (primaria, secundaria, universitaria, técnica y posgrado) reciban retornos diferentes. Este modelo es importante por dos razones: una es que Tenjo (1993) encuentra evi-

dencia para decir que los retornos a los distintos niveles educativos en Colombia no son homogéneos, y la segunda es que es posible que se encuentren diferencias importantes en las remuneraciones a los distintos niveles educativos en los dos sectores.

Además de las anteriores se incluyeron como variables dependientes dos variables para el tamaño de la planta y dos variables dummy para las características del lugar de residencia. Las variables para el tamaño de planta son importantes porque cambia la relación capital/trabajo y la probabilidad de que en la empresa haya un sindicato y estos dos factores pueden ser determinantes del salario. Estas variables son: tamaño de planta entre 11 y 49 trabajadores y tamaño de planta mayor a 49 trabajadores; el grupo de control es el de los trabajadores de empresas con menos de once empleados. Se incluyen como variables dummy porque así aparecen en la encuesta. Se puede pensar que en el sector público todas las empresas son grandes; de hecho en la ENII no se hace esta pregunta a los trabajadores del sector público y se los clasifica siempre como trabajando en empresas de más de once empleados. En la ENCV esta distinción es importante ya que la encuesta se realizó a nivel nacional (no como la ENH en la que la mayoría de sus etapas se hacen solamente en las ciudades principales) por lo que es posible encontrar trabajadores del sector público en empresas pequeñas.

Las variables cualitativas relacionadas con las características del lugar de residencia son: residencia en área rural o residencia en centro poblado; la característica de control es residencia en cabecera de municipio. Estas variables son importantes porque la encuesta se aplicó a nivel nacional.

En los estudios sobre los determinantes de los salarios en el sector público se incluyen variables adicionales diferentes al capital humano (Ehrenberg y Schwartz, 1986). Variables como las características del entorno político y el tamaño de la economía pueden ser importantes para determinar el salario, además del capital humano. Respecto de esto se pueden dar dos argumentos a favor de suponer que los salarios en el sector público se determinan como lo suponen las ecuaciones de Mincer. El primero es un argumento pragmático, aunque la especificación del modelo desarrollado permite que en cada una de las dos ecuaciones –(1) y (2)– se incluyan variables diferentes. Una de las cosas que se quiere captar en la investigación es cuáles son las diferencias en las remunera-

ciones al capital humano en los dos sectores, para ello necesitamos suponer que las variables que determinan los salarios son las mismas en los dos sectores (como se ha hecho en el modelo desarrollado). Además una de las técnicas de descomposición del diferencial requiere que las variables en las dos ecuaciones sean las mismas. El segundo argumento es que la mayoría de las variables que se reconocen como determinantes del salario en el sector público sirven para estudiar diferencias entre grandes regiones, en especial entre países, pero esto está fuera del propósito de la investigación.

Uno de los problemas importantes en este tipo de estudios es la construcción de la variable con el ingreso laboral. Es especialmente importante porque como dice Smith (1977) una parte importante de la compensación que se da a los empleados por su trabajo puede no estar incluida en el salario mensual. Hay otro tipo de pagos con una periodicidad diferente o que no se dan en dinero, que generalmente no se incluyen en los reportes que hacen los trabajadores de sus salarios; en general los trabajadores subreportan sus ingresos y en parte esto se debe a que sólo reportan los pagos que se dan con cierta periodicidad. La ENCV permite construir una mejor medida del salario con respecto a la que permite la ENH. En la ENCV se reportan, además del salario devengado periódicamente, otros pagos como subsidios de alimentación, transporte, vivienda, pagos en especie, primas anuales, y se reportan también los descuentos que se le hacen a los trabajadores por seguridad social y por retención en la fuente. Todas estas variables, además del pago periódico que reciben los trabajadores, se usaron en la construcción de la variable salario por hora que se utilizará en la estimación de los modelos que se expusieron en la sección anterior. De todas formas la encuesta no permite incluir como parte del salario otro tipo de beneficios no pecuniarios que pueden ser importantes.

La construcción de las variables en general no tiene problemas: tanto las variables dummy para el sexo, el tipo de educación más alto alcanzado por los individuos como el número de años, la característica del lugar de residencia y el tamaño de planta están disponibles en la ENCV. La variable que presenta problemas es la experiencia pues no está disponible en la ENCV por lo que se usó como proxy la experiencia potencial que es igual a la edad de la persona menos el número de años de educación menos cinco. Esta variable puede acercarse bastante a la experiencia real para el caso de los hombres pero no para el de las mujeres que han tenido tasas

de participación variables y menores a las de los hombres por lo que es posible que tengan discontinuidades en su experiencia. En muchos casos, para las estimaciones con muestras de mujeres se incluye, además de la experiencia, el número de hijos que cada mujer ha tenido, pero eso no es posible en este caso porque la encuesta no incluye esta información. Este es un punto que hace, además del sesgo introducido por la decisión de participar o no, menos confiables las estimaciones para las mujeres.

Respecto de las variables que se deben incluir en la ecuación de selección, la literatura no es clara. Sin embargo, se ve en el modelo de la tercera sección que en la ecuación de selección se deben incluir las mismas variables que en las ecuaciones de ingresos ya que parte del problema de selección está determinado por el diferencial entre los ingresos en los dos sectores. Se incluyeron en la ecuación de selección todas las variables que se consideraron en la ecuación de ingreso, excepto las variables dummy para el tamaño de planta.

No hay claridad sobre cuáles son las variables que determinan los costos incluidos por medio de la ecuación (4). Los costos asociados a la decisión pueden estar relacionados con la información disponible para el trabajador y con las preferencias de los trabajadores respecto del trabajo en cada uno de los sectores. Éstas se incluyen en este grupo de variables relacionadas con el entorno de la persona, confiando en que ellas den la información que se requiere. Éstas comprenden variables cualitativas como la posición ocupacional y educación del padre, y una variable dummy con características del tamaño del lugar en que la persona nació (que ya se mencionó en la sección anterior). También se incluyeron variables asociadas con las características de la persona: si es casado, si es jefe de hogar y si es migrante (esta última definida como que la persona nació en otro lugar y lleva viviendo en el área en que lo encuestaron menos de seis años).

Se reconoce a menudo que es posible que al utilizar este tipo de técnicas exista un problema de identificación. Sin embargo, Heckman y Honoré (1990) y Maddala (1983) muestran que cuando se trabaja con el supuesto de que los errores de las ecuaciones (1) y (2) tienen distribución normal, y por lo tanto se aplica el modelo probit para estimar los  $\lambda$ , al ser esta técnica no lineal, se evita el problema de identificación. De todas formas se recomienda que las dos ecuaciones (ecuación de salario y ecuación de selección) no tengan las mismas variables independientes. Algunos di-



cen que basta con que en la ecuación de selección haya al menos una variable adicional o diferente a las que se incluyen en la ecuación de ingresos (Asaad, 1997). Otros sugieren que es necesario que en cada ecuación haya una variable diferente a las de la otra ecuación (Dustmann y Van Soest, 1998). Siguiendo esto último, lo que se hizo fue incluir en la ecuación de selección, información respecto de la posición ocupacional, la educación del padre de los trabajadores y otras variables con las características de los trabajadores, y no incluir las dos variables cualitativas con el tamaño de planta.

Un punto importante para la estimación que se hizo es el del proceso de autoselección. La exposición que se hizo en la tercera parte del documento supuso que los trabajadores sólo tienen dos opciones entre las cuales escoger: trabajo en el sector público o trabajo en el sector privado. Como se dijo en esa misma sección, en realidad los trabajadores escogen entre un espectro más amplio de opciones. Es posible suponer muchas estructuras diferentes para este proceso aumentando el espectro de posibilidades que tiene el trabajador. Una de ellas es suponer que los trabajadores sólo tienen las dos opciones mencionadas e ignorar las demás. Otra es tener en cuenta el trabajo en otros sectores e incluir, por ejemplo, la decisión de ser o no trabajador asalariado. Una tercera posibilidad surge al tener en cuenta el desempleo como una de las opciones. La cuarta surge de si se tiene en cuenta la posibilidad de no participar en el mercado laboral.

El problema se complica aún más si se tiene en cuenta que es posible que muchas de estas decisiones son previas a la escogencia del sector en que se quiere trabajar. Por ejemplo, antes de decidir si se busca trabajo en el sector público o en el sector privado puede estar la decisión de si se es asalariado o independiente; anterior a ésta, está la decisión relativa al desempleo y, anterior a ésta, la que surge con la decisión de participación. De todas, la que exige un tratamiento independiente (si se decide tenerla en cuenta) es la de participación. Respecto de la decisión de ser asalariado o independiente se puede pensar que es simultánea con la decisión entre el sector público y el privado. En el caso del desempleo, en un marco de búsqueda de empleo, se puede pensar que una de las posibilidades de los agentes es permanecer desempleado frente a trabajar en alguno de los sectores. De alguna manera los agentes deciden permanecer desempleados si la utilidad de estar desempleado es mayor a la utilidad

en los trabajos que le ofrecen. En el caso de la decisión de participación, la decisión es anterior incluso a la de buscar trabajo, por lo que difícilmente se puede afirmar que la no participación es una opción que está al mismo nivel que la escogencia de sector.

En la estimación de este modelo se optó por una alternativa simple. El cálculo de los  $\lambda$  se hizo usando a todos los trabajadores empleados. No se tuvo en cuenta la decisión de participación, ni la relativa al desempleo. La primera se omite por la necesidad de incluir un nuevo proceso de selección que complica mucho la estimación, en especial la matriz de varianzas y covarianzas de los errores. La segunda se omite porque, aunque de alguna manera los agentes deciden permanecer desempleados, es imposible ignorar que el problema del desempleo está relacionado con factores de desequilibrio y problemas relativos a la duración y al ciclo de vida. Versiones preliminares de las estimaciones aquí presentadas tuvieron en cuenta procesos de autoselección más complejos pero los resultados de éstas no eran mejores de los que acá se presentan.

Sin embargo, sólo se hacen estimaciones de las ecuaciones de ingresos para los empleados asalariados del sector privado y para los empleados del sector público. La razón de esto es que en los reportes de los ingresos de los trabajadores no asalariados no es posible distinguir la parte correspondiente a la remuneración al capital humano y la remuneración a otros tipos de capital, lo que hace que la estimación de una ecuación de ingresos para los trabajadores de este sector no sea comparable con la de los trabajadores asalariados.

Como se mencionó anteriormente, se hicieron estimaciones de los modelos planteados en las ecuaciones (10) y (11) con la corrección del sesgo muestral apropiada que se explicó en la sección 2, según el resultado expuesto en la ecuación (14), y se usó toda la muestra de empleados para el cálculo de las probabilidades de estar en alguno de los dos sectores. Para tomar la decisión de cuál de los dos modelos es más apropiado, el modelo con *spline* y el modelo sin *spline*, se hizo la prueba F. En todos los casos en que se hizo la estimación se concluye que es más apropiado el modelo con *spline*, por lo que de ahora en adelante el documento se centra en él.

El cuadro 8 muestra los resultados de la estimación de la ecuación de selección del modelo con *spline*; los resultados son diferentes a los que se muestran en el cuadro 7 ya que en éstos se usa la muestra de todos los

trabajadores mientras que en los del Cuadro 8 sólo se utiliza la muestra de trabajadores asalariados. Además, en el Cuadro 8 se reportan los coeficientes del modelo probit y no las derivadas de la probabilidad de trabajar en el sector público o privado con respecto a cada variable, por lo que es difícil concluir algo acerca de la magnitud del efecto de cada variable sobre la probabilidad de trabajar en cualquiera de los dos sectores.

En general, para las seis estimaciones (hombres y mujeres separados y juntos, y sector público y sector privado) se encuentra que las variables que son siempre significativas son algunas de las que se incluyen en la ecuación de ingresos (edad, edad al cuadrado, años de educación, y sexo). Esto es evidencia que apoya el argumento de que lo que es realmente importante para la decisión de en qué sector se quiere trabajar es la diferencia entre los salarios que se obtendrían en cada uno de ellos.

Además de estas variables, también son significativas las variables para las características del lugar de residencia: esto se puede deber a diferencias entre los tres tipos de zonas en cuanto a la presencia o ausencia de entidades oficiales en ellas. Las demás variables no son siempre significativas, sin embargo son importantes la variable dummy para la educación universitaria del padre, que afecta positivamente la probabilidad de trabajar en el sector privado, y el ser jefe de hogar, que afecta positivamente la probabilidad de trabajar en el sector público.

La observación de los resultados puede sugerir la existencia de una incoherencia en ellos que es importante aclarar. Algunas variables, como por ejemplo la educación, tienen efecto positivo sobre la probabilidad de estar en cualquiera de los dos sectores. Si la base de datos utilizada sólo incluyera a los trabajadores asalariados, divididos entre trabajadores del sector público y trabajadores del sector privado, lo que se debería encontrar es que una variable que afecta positivamente la probabilidad de estar en un sector afecte negativamente la probabilidad de estar en el otro. Esto es lo que se ve en los resultados que se exponen en el Cuadro 7 y que fueron explicados en la sección anterior. En este caso la estimación de los  $\lambda$  se hizo mediante un probit teniendo en cuenta toda la muestra de empleados, por lo que es posible que se dé el resultado encontrado. Lo que el Cuadro 8 muestra no es el efecto de una variable sobre la probabilidad de estar en uno de los sectores con respecto al otro, sino la probabilidad de estar en uno de los sectores con respecto a la de estar en cualquiera de las otras dos posibilidades que hay en el mercado laboral.

**Cuadro 8** Modelo de Roy  
Ecuación de selección

	Sin separar sexos				Hombres		Mujeres	
	Sector público	Sector privado	Sector público	Sector privado	Sector público	Sector privado	Sector público	Sector privado
Dummy: mujer	0,4226 (8,39) *	-0,0169 (0,54)						
Dummy: lugar de nacimiento	0,0642 (1,20)	-0,1174 (3,71) *	0,0743 (1,03)	-0,1372 (3,47) *	0,0655 (0,80)	-0,0554 (0,97)		
Años de educación	0,1733 (11,77) *	0,0468 (5,29) *	0,1465 (7,52) *	0,0479 (4,54) *	0,2060 (8,86) *	0,0523 (3,10) *		
Años de educación secundaria	-0,0160 (0,94)	0,0131 (1,12)	0,0225 (0,98)	-0,0048 (0,34)	-0,0613 (2,34) *	0,0337 (1,61)		
Años de educación universitaria	-0,0738 (2,12) *	-0,0741 (3,13) *	-0,0440 (0,94)	-0,0793 (2,71) *	0,1079 (2,01) *	-0,0727 (1,70) *		
Años de educación técnica	0,1577 (3,02) *	0,0079 (0,21)	-0,1114 (1,58)	-0,0564 (1,20)	-0,2177 (2,77) *	0,0671 (1,05)		
Años de educación de posgrado	-0,0966 (0,98)	-0,3800 (4,76) *	-0,0241 (0,19)	-0,3041 (3,21) *	-0,1885 (1,19)	-0,5519 (3,56) *		
Años de experiencia potencial	0,0272 (4,36) *	-0,0129 (3,53) *	0,0226 (2,63) *	-0,0171 (3,72) *	0,0318 (3,31) *	-0,0041 (0,62)		
Años de experiencia potencial al cuadrado	0,0004 (3,37) *	0,00004 (0,77)	-0,0003 (2,18) *	0,00001 (0,18)	-0,0004 (2,25) *	-0,0002 (1,53)		
Dummy: padre con educación primaria	0,0778 (1,51)	-0,0221 (0,75)	0,0743 (1,07)	0,0734 (2,02) *	0,0787 (0,98)	0,0866 (1,57)		
Dummy: padre con educación secundaria	0,0196 (0,28)	0,0572 (1,24)	-0,0534 (0,55)	0,0119 (0,19)	0,0196 (0,18)	0,1835 (2,36) *		

(Continúa)

(Continuación Cuadro 8)

	Sin separar sexos		Hombres		Mujeres	
	Sector público	Sector privado	Sector público	Sector privado	Sector público	Sector privado
Dummy: padre con educación universitaria	-0.3693 (3.18)*	0.1008 (1.24)	0.3794 (2.41)*	0.0539 (0.50)	0.3326 (1.85)*	0.3793 (2.72)*
Dummy: padre obrero o empleado	0.0352 (0.65)	0.2898 (8.43)*	0.1168 (1.57)	0.3143 (7.22)*	-0.0814 (0.96)	0.2523 (4.24)*
Dummy: padre independiente o cuenta propia	0.0267 (0.49)	0.0173 (0.50)	0.0569 (0.75)	-0.0052 (0.12)	-0.0232 (0.28)	0.0469 (0.77)
Dummy: padre patrón o empleador	-0.1062 (0.82)	0.0150 (0.18)	-0.0595 (0.32)	0.0105 (0.09)	-0.1410 (0.74)	0.0180 (0.12)
Dummy: padre rentista	0.1627 (0.80)	-0.9639 (7.69)*	0.1809 (0.71)	-1.1559 (7.31)*	-0.0553 (0.16)	-0.3377 (1.33)
Dummy: casado o en unión libre	0.0756 (1.63)	0.0278 (0.94)	-0.0085 (0.10)	0.4153 (2.37)*	0.1575 (2.14)*	0.0624 (1.15)
Dummy: jefe de hogar	0.2399 (4.59)*	0.0483 (1.49)	0.2771 (2.99)*	-0.0200 (0.39)	0.3299 (3.85)*	0.033 (0.53)
Dummy: migrante	0.0254 (0.42)	0.1035 (2.83)*	0.0970 (1.23)	0.2257 (4.94)*	-0.0888 (0.90)	0.1290 (1.93)*
Dummy: residente de centro poblado	0.4132 (6.49)*	-0.4446 (10.10)*	0.2874 (3.39)*	-0.4092 (7.80)*	0.5629 (5.71)*	-0.5379 (6.11)*
Dummy: residente área rural dispersa	-0.0088 (0.12)	-0.6513 (16.27)*	-0.1731 (1.77)*	-0.6716 (14.55)*	0.2511 (2.01)*	-0.6385 (7.19)*
Constante	3.7662 (28.18)*	0.3852 (5.15)*	-3.5254 (20.17)*	0.2890 (3.35)*	3.6507 (18.08)*	0.6214 (4.91)*
Observaciones	12522	12522	8761	8761	3758	3758

Valores absolutos de los estadísticos z entre paréntesis.

\* significativo al 5%.

Fuente: cálculos del autor con datos de la Encuesta Nacional de Calidad de Vida, DANE, Colombia, 1997.

En el Cuadro 9 se reportan los resultados de la estimación de la ecuación de ingresos. En todas las ecuaciones estimadas las variables que tradicionalmente se han reconocido como determinantes del salario son significativas: la educación, la experiencia y el sexo. Los coeficientes de este cuadro se deben interpretar como tasas de cambio del salario con respecto a la variable correspondiente, cuando se trata de variables continuas, y como el porcentaje adicional que tiene el salario de una persona con determinada característica con respecto a una persona sin esa característica, cuando se trata de variables dummy.

Es importante observar en estas ecuaciones que el coeficiente asociado al  $\lambda$  es significativo en las estimaciones sin separar sexos y en las estimaciones para los hombres. Atendiendo a los resultados expuestos en las ecuaciones (10) y (11), se observa que el signo de la covarianza entre el error de la ecuación de ingresos y el de la ecuación de selección para los empleados del sector público es positivo y para los del sector privado es negativo. Estamos entonces en el caso en que los trabajadores de los dos sectores obtienen salarios promedio inferiores al promedio incondicional de su sector. Es importante recordar aquí que en este escenario los trabajadores escogen pertenecer a uno de por lo menos tres sectores (público, privado y no asalariado o independiente), por lo que es necesario entender este resultado en términos del aprovechamiento de ventajas competitivas, incluyendo a un tercer sector. Es por esto que es posible que los trabajadores de los dos sectores para los que se estiman ecuaciones de ingreso obtengan salario por debajo de los salarios medios incondicionales de sus sectores. Seguramente lo que está sucediendo aquí es que los trabajadores no asalariados son los que tienen mejores ventajas comparativas y podrían obtener los salarios más altos en cada uno de los tres sectores.

Aunque la observación de estas estimaciones no es suficiente para afirmar que existe un diferencial salarial entre los dos sectores, es posible ver en éstas que las estructuras de las remuneraciones son diferentes. Se encuentran tres diferencias en la significancia de algunas variables entre las ecuaciones de los dos sectores.

La primera de estas diferencias es que en el sector público los retornos a los distintos niveles educativos son iguales mientras que en el sector privado son diferentes. Las variables que hacen parte del *spline* de la educación cuando son significativas sólo lo son en las ecuaciones del sector privado. Específicamente se encuentra que en el sector privado el retor-

**Cuadro 9** Modelo de Roy  
Ecuación de salario por hora

	Sin separar sexos				Hombres		Mujeres	
	Sector público	Sector privado	Sector público	Sector privado	Sector público	Sector privado	Sector público	Sector privado
Dummy: mujer	-0.1365 (3.15)*	-0.1017 (4.36)*						
Años de educación	0.0734 (4.81)*	0.0838 (9.89)*	0.0536 (2.31)*	0.0802 (7.14)*	0.1197 (4.61)*	0.1065 (6.30)*		
Años de educación secundaria	-0.0192 (1.31)	-0.0265 (2.64)*	-0.0309 (1.40)	-0.0236 (1.77)*	-0.0088 (0.42)	-0.0257 (1.52)		
Años de educación universitaria	0.0414 (1.42)	0.0837 (3.96)*	0.0699 (1.56)	0.0836 (2.94)*	0.0007 (0.02)	0.0604 (1.64)		
Años de educación técnica	0.0414 (0.91)	0.0074 (0.25)	-0.0278 (0.39)	0.0095 (0.23)	-0.0863 (1.35)	0.0072 (0.15)		
Años de educación de posgrado	0.0007 (0.01)	0.3309 (4.71)*	0.0512 (0.45)	0.2493 (2.88)*	0.0792 (0.69)	0.4941 (3.16)*		
Años de experiencia potencial	0.0216 (3.17)*	0.0370 (12.39)*	0.0495 (4.84)*	0.0392 (9.83)*	-0.0046 (0.43)	0.0347 (6.81)*		
Años de experiencia potencial al cuadrado	0.0002 (1.51)	-0.0004 (6.81)*	-0.0008 (4.22)*	-0.0004 (6.04)*	0.0005 (2.15)*	-0.0004 (3.51)*		
Dummy: tamaño de empresa entre 11 y 49 empleados	-0.0026 (0.01)	0.1793 (6.41)*	0.6746 (1.03)	0.1273 (3.24)*	-0.4733 (0.68)	0.2519 (5.79)*		
Dummy: tamaño de empresa mayor a 49 empleados	0.3153 (0.81)	0.3843 (15.20)*	1.1058 (1.89)*	0.3662 (10.71)*	-0.3173 (0.55)	0.4183 (9.99)*		

(Continúa)

Diferenciales salariales entre los trabajadores  
del sector público y del sector privado  
en Colombia en 1997

Dario Maldonado Carrizosa

(Continuación Cuadro 9)

	Sin separar sexos		Hombres		Mujeres	
	Sector público	Sector privado	Sector público	Sector privado	Sector público	Sector privado
Dummy: residente de centro poblado	-0.1758 (2.93) *	0.0094 (0.21)	0.1588 (1.77) *	0.0287 (0.48)	0.1208 (1.27)	0.0786 (0.78)
Dummy: residente área rural dispersa	-0.1318 (1.63)	0.1115 (2.38) *	-0.0624 (0.52)	0.0492 (0.74)	-0.1554 (1.28)	-0.0588 (0.46)
Constante	6.9793 (15.45) *	6.1212 (76.63) *	6.0892 (8.68) *	6.0643 (51.06) *	6.9564 (9.68) *	5.6632 (24.60) *
$\lambda$	0.2888 (5.31) *	0.3011 (6.07) *	0.249 (2.70) *	0.2069 (2.34) *	0.1531 (1.29)	0.1127 (0.68)
$\rho$	0.4771	0.4463	0.4020	0.2964	0.2643	0.1792
$\sigma$	0.6053	0.6757	0.6194	0.6980	0.5792	0.6289
LR	-2902.35	-9868.61	-1542.95	-6472.74	-1315.76	3323.21
Test F (ret. hom. vs. ret. no hom.)	20.92 *	190.02 *	33.78 *	92.52 *	16.70 *	135.94 *

Valores absolutos de los estadísticos t entre paréntesis.

\* significativo al 5%.

Fuente: cálculos del autor con datos de la Encuesta Nacional de Calidad de Vida, DANE, Colombia, 1997.



no a la educación de posgrado es considerablemente mayor al retorno a los años de los demás niveles educativos.

La segunda diferencia que se encuentra es que el tamaño de planta no afecta significativamente el salario en el sector público pero sí lo afecta en el sector privado. Las dos variables asociadas a esta característica son significativas en las estimaciones para el sector privado y no lo son en las ecuaciones para el sector público. Parte de la explicación de esto se puede deber a la baja cantidad de empresas del sector público con tamaño inferior a 50 empleados y a que en el sector público las externalidades de las negociaciones en determinadas empresas son mayores que en el sector privado.

Además de lo anterior es importante ver las diferencias en las magnitudes de los coeficientes estimados, aunque no sea posible hacer pruebas de hipótesis sobre su igualdad ya que no se conocen las covarianzas de coeficientes de ecuaciones diferentes. No es posible a partir de un ejercicio como este saber en cuál de los dos sectores la remuneración es mayor; la razón está en que la evidencia es ambigua: se encuentra que en general las tasas de retorno al capital humano son mayores en el sector privado pero la constante es mayor en el sector público.

## **6. El diferencial salarial**

En esta sección se analiza la magnitud del diferencial salarial entre los trabajadores del sector público y los trabajadores del sector privado. Se usan dos tipos de medidas del diferencial salarial: la primera es la descomposición del diferencial promedio observado desarrollada en Oaxaca (1973) para el caso de la discriminación salarial por sexos; la segunda es un promedio de los diferenciales estimados para distintos niveles educativos, similar a la que usan Dustman y Van Soest (1998).

### **a. Descomposición de la diferencia entre los promedios de los logaritmos naturales de los salarios en los dos sectores**

La descomposición de Oaxaca se usa para calcular qué parte de la diferencia entre los promedios de los logaritmos de los salarios observados en

cada uno de los dos sectores se debe a diferencias en el capital humano poseído por los trabajadores de los dos sectores y qué parte se debe a diferencias en las remuneraciones al capital humano en los dos sectores. La descomposición se apoya en que por las propiedades de los Mínimos Cuadrados Ordinarios se tiene que:

$$\overline{\ln w_{0i}} = \hat{\beta}_0' \overline{X_{0i}^1} \quad \text{y} \quad \overline{\ln w_{1i}} = \hat{\beta}_1' \overline{X_{1i}^1} \quad (15)$$

y es como sigue

$$\begin{aligned} \overline{\ln w_{0i}} - \overline{\ln w_{1i}} &= \hat{\beta}_0' \overline{X_{0i}^1} - \hat{\beta}_1' \overline{X_{1i}^1} = \hat{\beta}_0' \overline{X_{0i}^1} - \hat{\beta}_1' \overline{X_{0i}^1} + \hat{\beta}_1' \overline{X_{0i}^1} - \hat{\beta}_1' \overline{X_{1i}^1} \\ &\Leftrightarrow \overline{\ln w_{0i}} - \overline{\ln w_{1i}} = \Delta \hat{\beta}' \overline{X_{0i}^1} + \hat{\beta}' \Delta \overline{X_{1i}^1} \end{aligned} \quad (16)$$

En estas expresiones  $\overline{X_{0i}^1}$  y  $\overline{X_{1i}^1}$  son matrices con los promedios de la matriz  $X_{ij}^1$  para los sectores 0 y 1 respectivamente. Además se ha incluido en los vectores de parámetros y en las matrices de la información, los componentes referentes a la corrección del sesgo de selección.

En la última parte de la expresión en (16) se muestra la descomposición. La primera parte es el diferencial salarial promedio debido a diferencias en los parámetros, evaluado en el promedio de las características de capital humano de uno de los dos sectores. La segunda parte es la porción del diferencial debido a diferencias en el capital humano, evaluado según los parámetros estimados para el sector privado. En esta expresión se evalúa el diferencial por parámetros en las características de los empleados del sector público y el diferencial por capital humano en los parámetros del sector privado. Es posible obtener una descomposición análoga en la que se evalúen las magnitudes del diferencial en los parámetros del sector público y en las características de los empleados del sector privado.

Los resultados de la descomposición de Oaxaca aparecen el Cuadro 10. En la primera columna, en la parte del cuadro en que se muestra la descomposición, ésta se hace según la ecuación (16), la segunda columna de la descomposición se hace según la otra posibilidad. De acuerdo con estos resultados se encuentra que el diferencial salarial entre los trabajadores del sector público y del sector privado, debido a diferencias en la remuneración al capital humano en los dos sectores, es positivo y de una

**Cuadro 10** Descomposición del diferencial salarial promedio

	Sin separar sexos		Hombres		Mujeres	
	Sector público (1)	Sector privado (2)	Sector público (1)	Sector privado (2)	Sector público (1)	Sector privado (2)
Promedio del logaritmo del salario estimado: parámetros sector público	8,3826 (0,0138)	7,931 (0,0066)	8,3499 (0,0221)	7,5711 (0,0142)	8,152 (0,0222)	7,9525 (0,0119)
Promedio del logaritmo del salario estimado: parámetros sector privado	8,1042 (0,0203)	7,5201 (0,0084)	7,9914 (0,0278)	7,4349 (0,0104)	7,9738 (0,0342)	7,3251 (0,0150)
Diferencial total	0,8625		0,915		0,8269	
Diferencial por diferencias en los parámetros	0,2784 32,28%	0,4299 48,80%	0,3585 39,18%	0,1362 14,89%	0,1782 21,55%	0,6274 75,87%
Diferencial por diferencias en el capital humano	0,5841 67,72%	0,4416 51,20%	0,5565 60,82%	0,7788 85,11%	0,6487 78,45%	0,1995 24,3%

Errores estándar entre paréntesis.

- (1) Diferenciales por parámetros evaluados en el capital humano del sector público y diferenciales por capital humano evaluados en los parámetros del sector privado.  
(2) Diferenciales por parámetros evaluados en el capital humano del sector privado y diferenciales por capital humano evaluados en los parámetros del sector público.

Fuente: cálculos del autor con datos de la Encuesta Nacional de Calidad de Vida, DANE, Colombia, 1997.

magnitud considerable, a favor de los trabajadores del sector público. En el primer caso, las diferencias en las remuneraciones al capital humano explican el 32.28% de la diferencia observada entre los promedios y las diferencias en el capital humano explican el 67.72% del diferencial promedio. En el caso de los hombres, las diferencias entre los parámetros de los dos sectores explican el 39.18% de la diferencia entre los promedios. En el caso de las mujeres, las diferencias por parámetros explican el 21.55% del diferencial promedio total. Llama la atención la gran diferencia que se encuentra entre las dos formas de descomponer el diferencial, en especial para el caso de las mujeres.

## b. Diferenciales promedio estimados

La segunda medida del diferencial es más simple y consiste en encontrar para cada trabajador el salario estimado en los dos sectores. Luego con estas dos estimaciones se obtiene el diferencial estimado para cada trabajador. La medida consiste en promediar el diferencial para distintos grupos de trabajadores. Matemáticamente esta medida se expresa de la siguiente manera:

$$\text{Diferencial estimado promedio} = \frac{\hat{w}_{0i} - \hat{w}_{1i}}{\hat{w}_{1i}} \quad (17)$$

en donde la barra sobre las variables indica que se trata de un promedio.

La importancia de un análisis como este se debe a que, dado que todos los trabajadores pueden tener productividades marginales diferentes, el salario de referencia con el que se debe hacer la comparación es el salario contrafactual de cada trabajador, es decir el que ganaría en el otro sector. La estimación del modelo de Roy para este propósito es especialmente importante porque permite incluir en la estimación del salario contrafactual las ventajas comparativas de cada trabajador. El cálculo de esta medida del diferencial, utilizando un modelo en el que el proceso de selección no se tiene en cuenta, no sería confiable al no incluir las características no observables de los trabajadores que se incluyen en esta estimación a través de la corrección del sesgo de selección.

El Cuadro 11 muestra los resultados de los diferenciales promedio estimados. En él se reportan estimaciones del diferencial extraídas de las estimaciones sin separar sexos, para los hombres y para las mujeres. En



Cuadro 11 Promedio de los diferenciales salariales estimados

	Sin separar sexo				Hombres				Mujeres			
	Diferencial incondicional	Diferencial condicional trabajadores del sector público	Diferencial condicional trabajadores del sector privado	Diferencial incondicional	Diferencial condicional trabajadores del sector público	Diferencial condicional trabajadores del sector privado	Diferencial incondicional	Diferencial condicional trabajadores del sector público	Diferencial condicional trabajadores del sector privado	Diferencial incondicional	Diferencial condicional trabajadores del sector público	Diferencial condicional trabajadores del sector privado
Total	51,34% (12,17)*	51,69% (125,61)*	35,60% (34,84)*	25,07% (32,40)*	21,68% (25,35)*	45,55% (35,43)*	87,45% (37,97)*	105,72% (38,56)*	25,37% (11,84)*			
Sin educación	46,96% (24,80)*	47,31% (24,58)*	34,38% (3,83)	5,21% (1,03)	4,9% (0,97)	25,06%*	148,79% (8,16)*	155,04% (7,90)*	89,36% (21,26)*			
Primaria	51,60% (79,57)*	52,53% (76,72)*	41,61% (25,33)*	24,39% (15,28)*	22,39% (13,55)*	53,34% (11,97)*	99,69% (13,83)*	109,75% (13,49)*	36,37% (6,74)*			
Secundaria	64,16% (12,14)*	66,00% (134,68)*	54,28% (53,91)*	38,79% (26,69)*	23,97% (20,19)*	56,82% (51,75)*	110,82% (38,09)*	124,04% (37,58)*	47,24% (32,00)*			
Técnico	34,00% (31,63)*	34,38% (27,62)*	32,60% (15,61)*	17,53% (5,66)*	13,68% (3,85)*	35,94% (11,35)*	52,57% (10,78)*	62,10% (10,35)*	22,66% (5,03)*			
Univ. sin título	32,85% (22,61)*	34,27% (19,76)*	27,72% (12,68)*	25,78% (7,37)*	22,82% (5,88)*	43,30% (10,79)*	46,69% (6,08)*	59,92% (5,64)*	18,24% (6,64)*			
Univ. con título	19,35% (30,95)*	20,60% (24,38)*	17,45% (19,48)*	25,44% (18,66)*	22,00% (11,40)*	31,81% (26,27)*	32,07% (10,66)*	47,33% (9,70)*	12,68% (9,23)*			
Posgrado	-19,48% (9,68)*	18,18% (6,03)*	20,39% (7,53)*	-7,67% (3,03)*	-10,53% (3,16)*	4,41% (1,35)	39,86% (11,31)*	32,95% (3,77)*	42,17% (11,94)*			

Valores absolutos de los estadísticos t entre paréntesis.

\* significativo al 5%.

Fuente: cálculos del autor con datos de la Encuesta Nacional de Calidad de Vida, DANE, Colombia, 1997.

cada caso se reportan tres tipos de estimaciones: en la primera columna aparece la estimación incondicional del diferencial, es decir se usa toda la muestra de trabajadores, en la segunda y tercera columnas aparecen estimaciones condicionales del diferencial. En la segunda se reportan los diferenciales estimados usando la muestra de trabajadores del sector público y en la tercera se reportan los resultados construidos usando la muestra de trabajadores del sector privado.

Los diferenciales promedio estimados son significativamente distintos de cero en casi todos los casos. En los resultados se ve que para casi todos los niveles educativos los salarios favorecen a los trabajadores del sector público. Para el caso de los trabajadores con nivel educativo de posgrado el resultado es el inverso, el salario promedio estimado de los trabajadores del sector privado es mayor que el de los del sector público. Esto es consistente con la gran diferencia que se encontró en la sección anterior en la remuneración a la educación de posgrado entre los sectores público y privado.

Puede parecer que las conclusiones de esta sección son contrarias a las de la quinta parte del documento. Allí se dijo que los salarios medios incondicionales de los dos sectores eran menores a los salarios medios incondicionales respectivos. Para entender por qué no hay contradicciones es importante tener en cuenta que en la conclusión de esa sección se hacía una comparación con referencia a toda la estructura de la remuneración al trabajo del mercado, mientras que la conclusión de esta sección sólo se refiere a la estructura de la remuneración a los trabajadores asalariados.

## 7. Conclusión

La conclusión básica de este documento es que hay evidencia para afirmar que existe un diferencial estadísticamente significativo entre los sectores público y privado, a favor de los empleados del sector público.

Las estructuras de las remuneraciones en los dos sectores son diferentes. El capital humano y otro tipo de determinantes del salario como el tamaño de planta tienen efectos diferentes sobre los salarios en los dos sectores. A partir del ejercicio de estimación de las ecuaciones mincerianas con retornos a los distintos tipos de educación no homogéneos se puede

ver que el retorno a la educación en el sector privado es mayor que en el sector público para la mayoría de los niveles educativos y que algunas de las variables que hacen parte del *spline* de educación son significativas solamente en las ecuaciones para el sector privado. Las dos variables categóricas para el tamaño de planta son significativas en las tres ecuaciones para el sector privado y sólo una de ellas es significativa en una de las ecuaciones para el sector público. En el mismo ejercicio se ve que el intercepto es mayor en las ecuaciones para el sector público.

Los diferenciales promedio estimados son significativos y su magnitud es muy variable, va desde un 42.47% a favor de las mujeres trabajadoras del sector privado con educación de posgrado a un 155% a favor de las mujeres trabajadoras del sector público sin educación. La descomposición de Oaxaca deja ver que alrededor del 30% del diferencial salarial entre los dos sectores se debe a diferencias en las estructuras de la remuneración al trabajo en los dos sectores.

En las estimaciones del modelo que se hicieron sin separar sexos y solamente con los hombres se rechazó la hipótesis nula de que el sesgo de selección no afecta los salarios en los dos sectores. La conclusión es que los trabajadores de los dos sectores obtienen salarios inferiores a los salarios medios incondicionales de su sector. Esto es posible ya que en el mercado laboral, además de los dos sectores para los que se hicieron las estimaciones, hay un sector de trabajadores no asalariados.

En el estudio de Perfetti (1999) se encuentran conclusiones que en cierto sentido son similares a las que se han encontrado en este estudio pero que también difieren de ellas. En Perfetti (1999) se halla que el diferencial favorece a los trabajadores del sector público. Sin embargo, en este estudio se ha encontrado que esto es cierto salvo para los trabajadores con educación de posgrado para los cuales es mayor el salario por hora promedio en el sector privado. En este estudio no se han estimado primas salariales como las de Perfetti (1999), pero éstas se pueden comparar con los promedios de los diferenciales estimados de este estudio, aunque en ellas no se aísla el efecto de otras variables que afectan el salario. En el estudio de Perfetti la prima del sector público para distintos niveles educativos es mucho más estable que la que se ha encontrado en este estudio: en aquel la prima es del orden de 15% para los empleados con educación primaria y secundaria, de 56% para los empleados sin edu-

cación, y de 7% para los empleados con educación superior. En cambio, la dispersión encontrada en este estudio respecto de las primas salariales entre los trabajadores de los dos sectores es mucho mayor, y las magnitudes son diferentes; en general se ve que el promedio de los diferenciales salariales estimados para grupos con distinto nivel educativo es mayor que las primas encontradas por Perfetti (1999).

Para terminar, cabe hacer algunas recomendaciones en términos de información que sería importante tener en cuenta para estudios similares a éste. Como se ha mencionado, una parte importante del diferencial puede deberse a diferencias en las condiciones laborales de los dos sectores. Una posible hipótesis para explicar el diferencial que se ha encontrado, que no se ha podido contrastar en este estudio, es que los trabajadores del sector público tienen condiciones laborales menos favorables que los trabajadores del sector privado (aunque las estadísticas descriptivas de la cuarta parte del trabajo permiten hacer la afirmación contraria). Para contrastar esta hipótesis y poder calcular un diferencial que no esconda este tipo de condiciones laborales sería importante disponer de información referente a las características de los empleos, y esto no aparece en la ENCV. Sólo con una base de datos con este tipo de información se podría afirmar el origen del diferencial.

Es probable también que el diferencial se deba a habilidades específicas requeridas por cada sector y que no es posible captar en las variables que se han usado en este estudio. Valdría la pena buscar información que permita distinguir el tipo de tareas que hacen los empleados de cada uno de los dos sectores, de manera que se puedan estimar diferenciales que no escondan este tipo de características de los trabajadores.

Una de las limitaciones del trabajo es que no se hace la comparación de los salarios del sector público con los de los trabajadores no asalariados. Se puede pensar que las distorsiones en el mercado de trabajo introducidas entre los trabajadores no asalariados son menores por la movilidad de éstos y por el tamaño de las unidades en las que trabajan, por lo que sería interesante hacer la comparación con este grupo de trabajadores. Como se explicó, esto se debe a que las remuneraciones de los trabajadores no asalariados esconden ganancias de capital que no es posible separar de la remuneración al trabajo con la información de la ENCV. Otra posible forma de mejorar la información para un estudio de este tipo es, incluir



información que permita separar las ganancias por capital físico de las ganancias por capital humano en las remuneraciones a los trabajadores no asalariados.

## Bibliografía

- Asaad, R. (1997). "The Effects of Public Sector Hiring and Compensation Policies on the Egyptian Labor Market". *The World Bank Economic Review*, 11/1.
- Dustman, C. y A. van Soest (1998). "Public and Private Sector Wages of Male Workers in Germany". *European Economic Review*, 42.
- Ehrenberg, R. G. y J. L. Schwartz (1986). "Public Sector Labor Markets". *Handbook of Labor Economics*, II. Eds.: O. Ashenfelter, R. Layard.
- Gómez, Luis Carlos (1998). "Diseño e implementación de la muestra probabilística". Encuesta de Calidad de Vida 1997, DANE.
- Heckman, J. J. (1976). "The Common Structure of Statistical Model of Truncation and Limited Dependent Variables and a Simple Estimator for Such Models". *Annals of Economic and Social Measurement*, 5/4.
- (1979). "Sample Selection a Specification Error". *Econometrica*, 47/1.
- y B. E. Honoré (1990). "The Empirical Content of the Roy Model". *Econometrica*, 58/5.
- Maddala, G. S. (1983). "Limited-Dependent and Qualitative Variables in Econometrics". *Econometric Society Monographs* No. 3. Cambridge University Press.
- Oaxaca, R. (1973). "Male-Female Wage Differentials in Urban Labor Markets". *International Economic Review*, 14/3.
- Perfetti del Corral, M. (1999). "Diferenciales salariales entre trabajadores asalariados hombres del sector público y del sector privado, en-

tre 1984 y 1994 en Colombia". en *Pobreza y desigualdad*. Eds.: M. Cárdenas, N. Lustig.

Psacharopoulos, G. (1983). "Education and Private versus Public Sector Pay". *Labour and Society*, 8/2.

Smith, S. (1977). *Equal Pay in the Public Sector: Fact or Fantasy*. Princeton University.

Tenjo, J. (1992). "Labour Markets, the Wage Gap and Gender Discrimination: The Case of Colombia", en *Women's Employment and Pay in Latin America*. World Bank. Ed. George Psacharopoulos.

— (1993). "Evolución de los retornos a la inversión en educación, 1976-1989". *Planeación y Desarrollo*, 24.