DESCOMPOSICIÓN OAXACA-BLINDER EN MODELOS LINEALES Y NO LINEALES

Marzo 2012 **nº 20**

José Vicéns Otero

El presente documento realiza una exposición del popular método de descomposición de una variable en dos grupos propuesto por Blinder (1973) y Oaxaca (1973). En su planteamiento original, la descomposición Blinder-Oaxaca determina que la diferencia de salarios existente entre dos grupos se debe a dos componentes. El primer de ellos recoge la diferencia entre las variables explicativas observables de los dos grupos y el segundo componente recoge la diferencia entre las características no observables, medido por las discrepancias entre los parámetros de ambos grupos. En el documento se exponen las alternativas de estimación tanto en modelos lineales como no lineales, los problemas que conlleva cada una de ellas y las diferentes alternativas de ponderación de los componentes de la descomposición. (*)

Palabras clave: Descomposición de diferencias, Blinder, Oaxaca, discriminación.

(*) Mi agradecimiento a las profesoras Eva Medina y Ainhoa Herrarte por sus críticas y correcciones al documento original.

Edita:

Instituto L.R.Klein – Centro Gauss
Facultad de CC.EE. y EE.
Universidad Autónoma de Madrid
28049 Madrid
Teléfono y Fax: 913974191
Correo Electrónico: klein.gauss@uam.es
Página Web: www.uam.es/klein/gauss

ISSN 1696-5035

Depósito Legal: M-30165-2003

1. PLANTEAMIENTO

El objetivo de este documento es presentar una aproximación al popular método de descomposición de una variable en dos grupos propuesto por Blinder (1973) y Oaxaca (1973) y conocido en la literatura como descomposición Oaxaca-Blinder.

En su planteamiento inicial, la descomposición Oaxaca-Blinder determina que la diferencia de salarios existente entre dos grupos de individuos se debe a dos componentes. El primero de ellos recoge la diferencia entre las variables explicativas observables de los dos grupos y el segundo recoge la diferencia entre las características no observables, medido éste por las discrepancias entre los parámetros de ambos grupos.

Entre los muchos desarrollos y extensiones que la literatura ha realizado sobre la descomposición Oaxaca-Blinder hay que destacar la utilización de la regresión cuantílica (Albrecht et al., 2003) (Mahadevan y Suardi, 2012) y las aplicaciones con datos binarios (Fairlie, 1999, 2005; Yun, 2004), existiendo multitud de aplicaciones a diferentes contextos en sus casi 40 años de existencia.

En general se supone la existencia de dos grupos de individuos A y B (hombres y mujeres, inmigrantes y nativos, etc.) con diferencias en los salarios percibidos por cada grupo en función de un conjunto K de variables explicativas o predictores *X*. El objetivo del análisis es determinar que parte de la diferencia en la variable endógena entre los dos grupos se debe a las diferencias existentes entre las variables explicativas y que parte a otros factores o variables.

En la econometría convencional el tratamiento de las diferencias suele realizarse mediante la incorporación de una variable ficticia, en la forma:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_k X_{ki} + \alpha G_i + \varepsilon_i$$

donde *G* es una variable ficticia que toma el valor 1 para un grupo y 0 para el otro. Esta alternativa considera la existencia de los grupos como explicación de los salarios y modifica el término constante para cada grupo. Sin embargo, no permite conocer la influencia diferenciada sobre la endógena de las variables explicativas en cada grupo si no se incluyen interacciones.

2. DESCOMPOSICIÓN DE OAXACA, BLINDER Y NEUMARK

Con la metodología que se expone a continuación sí va a ser posible diferenciar por grupos la influencia de los regresores. El trabajo inicial se debe a Blinder (1973) y Oaxaca (1973) y ha tenido una amplia aceptación en el contexto de la econometría aplicada. El objetivo es conocer cuánto de la diferencia existente en la endógena es consecuencia de las diferencias en los predictores por grupos y cuánto se debe a otros factores:

$$E(\,Y_A\,) - \,E(\,Y_B) = \,R$$

Donde $E(Y_A)$ es la esperanza matemática de la variable endógena Y para el grupo A, $E(Y_B)$ es la esperanza de Y en el grupo B

El modelo lineal general que consideramos será:

$$Y_g = \,X_g\beta_g + \varepsilon_g; \quad E\left(\varepsilon_g\right) = 0 \quad g\epsilon\{A,B\}$$

y con ello:

$$R = \,E(\,Y_A\,) - \,E(\,Y_B) = E(\,X_A\beta_A\,) - E(X_B\beta_B\,)$$

Si en lugar de las esperanzas matemáticas consideramos los valores medios tenemos que:

$$R = \,\bar{Y}_A - \bar{Y}_B$$

dónde $\bar{Y}_A$ y $\bar{Y}_B$ representan el valor medio de la variable Y para el grupo A y B respectivamente.

A partir de

$$Y_A = X_A\beta_A + \varepsilon_A$$

$$Y_B = X_B\beta_B + \varepsilon_B$$

tendremos

$$\bar{Y}_A = \bar{X}_A\hat{\beta}_A$$

$$\bar{Y}_B = \bar{X}_B\hat{\beta}_\beta$$

con lo que el estimador de las diferencias será:

$$R = \bar{Y}_A - \bar{Y}_B = \bar{X}_A\hat{\beta}_A - \bar{X}_B\hat{\beta}_B$$

Sumando y restando $\bar{X}_B\hat{\beta}_A$ en la expresión anterior

$$R = \bar{X}_A\hat{\beta}_A - \bar{X}_B\hat{\beta}_B + \bar{X}_B\hat{\beta}_A - \bar{X}_B\hat{\beta}_A$$

y sacando factor común

$$R = (\bar{X}_A - \bar{X}_B)\hat{\beta}_A + \bar{X}_B(\hat{\beta}_A - \hat{\beta}_B) \qquad [1]$$

obtenemos la ecuación básica de la descomposición. Según la ecuación [1] la diferencia en los salarios de los grupos A y B es igual a la suma de dos componentes. El primer componente se interpreta como la parte explicada por las diferencias entre las variables observadas y el segundo componente como la parte no explicada o perteneciente a variables no observadas.

Una forma alternativa de plantarlo es mediante una regresión global. Para ello introducimos una variable ficticia G_i de forma que:

G_i = 0 para los individuos pertenecientes al grupo de referencia
G_i = 1 para los individuos pertenecientes al grupo de comparación

Admitiendo que en uno o en otro grupo podríamos poner tanto al grupo A como al B, efectuamos la regresión

$$Y_i = \alpha_0 + \alpha_1 G_i + \varepsilon_i$$

considerando G_i =1 para el grupo A y G_i =0 para el grupo B, tendremos que:

$$E\ (Y_{Ai}) = \alpha_0 + \alpha_1$$

$$E\ (Y_{Bi}) = \alpha_0$$

y la diferencia

$$R = E\ (Y_{Ai}) - E\ (Y_{Bi}) = \alpha_1$$

Con lo que la descomposición de Oaxaca-Blinder tendría por objetivo explicar α_1, que es el parámetro que recoge la diferencia entre los grupos.

En la ecuación [1] y en estudios empíricos con salarios de grupos alternativos (hombres y mujeres; blancos y negros; etc.) a la parte no explicada suele considerársela como una medida de la discriminación. En el grupo que se considera discriminado, pongamos por ejemplo el B, podríamos estimar un valor de la endógena exento de esta discriminación por:

$$\bar{Y}_{0B} = \bar{X}_B\, \hat{\beta}_A$$

indicando el subíndice 0 la ausencia de discriminación.

En este caso la estructura de salarios del grupo A (por ejemplo hombres) sería la norma de comparación. De igual forma cuando se considera que el salario no discriminatorio es el del grupo B (mujeres), en lugar de la expresión [1] tendremos:

$$R = (\bar{X}_A - \bar{X}_B)\hat{\beta}_B + \bar{X}_A(\hat{\beta}_A - \hat{\beta}_B) \qquad [2]$$

Las ecuaciones [1] y [2] pueden plantearse alternativamente por:

$$R = \Delta\overline{\mathrm{X}}\, \hat{\beta}_A + \bar{X}_\beta \Delta\hat{\beta}$$

$$R = \Delta\overline{\mathrm{X}}\, \hat{\beta}_B + \bar{X}_A \Delta\hat{\beta}$$

y en ambas la primera parte recoge la diferencia entre las características de los dos grupos y la segunda las diferencias en los coeficientes o diferencias en la estructura de salarios. Inicialmente podría pensarse que la proporción que corresponde a las diferencias en las características y la proporción derivada de las diferencias en las estructuras deberían coincidir en ambas ecuaciones. Sin embargo este no es el caso y los resultados obtenidos en trabajos empíricos con el porcentaje debido a las características y el porcentaje debido a la estructura salarial o parte discriminatoria, difieren según se tome la ecuación [1] ó [2]. Así en el trabajo original de Oaxaca sobre el diferencial salarial de hombres y mujeres, se obtiene que un 53% de dicho diferencial es debido a la discriminación basándose en la ecuación [1], mientras que dicho porcentaje se eleva al 64% a partir de la ecuación [2].

Neumark (1988) critica el planteamiento inicial de Oaxaca de considerar este inconveniente como un simple problema de "números índices" y de elección de la estructura salarial no discriminatoria de referencia (hombres o mujeres) y plantea una descomposición general en la que exista una estructura teórica de salarios no discriminatorios, es decir un vector de parámetros $\hat{\beta}_o$ no discriminatorio tal que:

$$R = \Delta\bar{X}\,\hat{\beta}_o + \left[\bar{X}_A\left(\hat{\beta}_A - \hat{\beta}_o\right) - \bar{X}_B\left(\hat{\beta}_B - \hat{\beta}_o\right)\right] \qquad [3]$$

En la descomposición propuesta por Neumark el primer término de la ecuación [3] es la diferencia de salarios debida a las diferencias en las características de los grupos, es decir, la parte no discriminatoria. El segundo término es la parte debida a la discriminación y que puede darse tanto para el grupo A como para el grupo B. Obsérvese que si consideramos que el grupo A no es discriminatorio $(\hat{\beta}_o = \hat{\beta}_A)$ la expresión coincide con la [1]. Por el contrario si el grupo no discriminatorio es el B $(\hat{\beta}_o = \hat{\beta}_B)$ la expresión [3] coincide con la [2].

Estimar $\hat{\beta}_o$ requiere una estimación promedio de los datos disponibles y la alternativa propuesta por Neumark es similar a:

$$\hat{\beta}_0 = [P_A(X'_A X_A) + P_B(X'_B X_B)]^{-1}[P_A(X'_A Y_A) + P_B(X'_B Y_B)]$$

siendo P_A y P_B la proporción de elementos del grupo A y del grupo B sobre el total de la muestra.

En un trabajo posterior, Oaxaca y Ransom (1994) generalizan la utilización de los parámetros mezcla o no discriminatorios determinando su valor mediante la expresión

$$\hat{\beta}_0 = \Omega\hat{\beta}_A + (I - \Omega)\hat{\beta}_B$$

En esta expresión la propuesta de Oaxaca y Blinder se obtiene cuando **Ω** = I, pero también otras alternativas de la literatura pueden recogerse. La de Reimers (1983):

$$\Omega = 0{,}5\ I$$

y la de Cotton (1988):

$$\Omega = s\ I$$

siendo s la proporción del tamaño del grupo mayoritario representado por el grupo A.

Por su parte Oaxaca y Ransom se alinean con la propuesta de estimar los parámetros promedio y en general la literatura y posteriores aplicaciones han seleccionado mayoritariamente la alternativa de Neumark.

Sin embargo un trabajo relativamente reciente de Elder, Godderis y Haider (2010), relativiza y pone en duda las aplicaciones realizadas con la descomposición de Oaxaca-Blinder utilizando la regresión mezcla de Neumark. La crítica se basa en que la regresión sin discriminación de Neumark no incluye un término independiente que diferencie los dos grupos, diferenciación que de hecho sí existe en la realidad y que es el punto de partida de todos los análisis. Al realizar la regresión para obtener los parámetros no discriminatorios (β_0) y omitir la diferencia de grupos, se obtendrán unos parámetros sesgados por omisión de variables, y la consecuencia será que en [3] se tenderá a sobrevalorar la parte explicada y a infravalorar la parte no explicada o asociada a la discriminación. Por ello estos autores proponen realizar la regresión planteada al principio de este documento

$$Y_i = \beta_0 + \beta_1 X_{1i} + \cdots + \beta_k X_{ki} + \alpha G_i + \varepsilon_i \qquad [3']$$

en la que el parámetro α recogería la parte no explicada de la diferencia entre los grupos, en el caso de que los coeficientes β no varíen entre los grupos. En este caso la parte no explicada en [3] es menor que la parte no explicada en [3'] y este resultado se mantiene aún cuando los coeficientes varían. Con ello se pone en duda la conveniencia de aplicar la descomposición de Oaxaca-Blinder con la solución de Neumark.

3. Corrección del sesgo de selección

En muchos de los trabajos con la descomposición de Oaxaca-Blinder se considera que las estimaciones por mínimos cuadrados pueden estar sesgadas por el hecho de que haya un proceso de selección muestral o de clasificación. El ejemplo más habitual de sesgo de selección se presenta cuando en una ecuación de salarios las variables que explican el salario influyen en la decisión de trabajar o no y la muestra de individuos incluye solamente aquellos que trabajan. Otro ejemplo de sesgo de selección se daría cuando los grupos se definen por la pertenencia o no a un sindicato y las variables que influyen en los salarios fueran las mismas que determinan la decisión de sindicarse, sean observables o no. Desde un punto de vista econométrico el problema es que la estimación por MCO obtendrá estimadores inconsistentes y sesgados ya que la selección estará correlacionada con la variable endógena y con el término de error. Ante esta situación la solución generalmente adoptada en la literatura por su sencillez ha sido la propuesta por Heckman (1979) en un proceso de estimación en dos etapas.

En la primera etapa se estima el modelo de interés o ecuación de resultado y en la segunda se recoge el proceso de selección muestral.

Ecuación resultado:

$$y_i = \begin{cases} X_i\beta + \varepsilon_i & si\ Z_i^* > 0 \\ --- & si\ Z_i^* \leq 0 \end{cases}$$

Ecuación de selección:

$$Z_i^* = W_i\gamma + u_i$$

con

$$Z_i = \begin{cases} 1 & si\ Z_i^* > 0 \\ 0 & si\ Z_i^* < 0 \end{cases}$$

donde W_i es un vector de variables observables y la ecuación se estima mediante un modelo probit.

Si las variables W_i son independientes no habrá problemas con la estimación de la ecuación resultado, pero sí se presentan problemas cuando los términos de error u_i y ε_i están correlacionados. En ese caso tendremos:

$$u_i \sim N\ (0{,}1)$$

$$\varepsilon_i \sim N\ (0, \sigma^2)$$

$$Corr\ (u_i\,, \varepsilon_i) = \rho$$

y para los datos observados

$$E(y_i/y_i\ observada)\ = E\ (X_i\beta + \varepsilon_i/W_i\gamma + u_i\ > 0)$$

$$= X_i\beta + E\ (\varepsilon_i/W_i\gamma + u_i\ > 0)$$

$$= (X_i\beta + E\ (\varepsilon_i/(u_i\ > -W_i\gamma)$$

En el caso de que ε_i y u_i sean independientes $E(\varepsilon_i) = 0$, y

$$E(y_i/y_i\ observada)\ =\ X_i\beta$$

puede estimarse por MCO. Pero si hay correlación entre los dos errores necesitamos calcular $E\ (\varepsilon_i/(u_i\ > -w_i\gamma)$

Para ello definimos:

$$\alpha_u\ =\ \frac{-w_i\gamma}{\sigma_u}$$

y el ratio λ conocido como el Ratio inverso de Mills

$$\lambda_{\,i}(\alpha_u) = \frac{\emptyset\left(\frac{-w_i\gamma}{\sigma_u}\right)}{1-\Phi\left(\frac{w_i\gamma}{\sigma_u}\right)} = \frac{\emptyset\left(\frac{w_i\gamma}{\sigma_u}\right)}{\Phi\left(\frac{w_i\gamma}{\sigma_u}\right)}$$

Ahora, tal y como se recoge en Greene (2003) tendremos que

$$E\ (\varepsilon_i/u_i \geq w_i\gamma) = \ \rho\sigma_\varepsilon \left[\frac{\emptyset\left(\frac{w_i\gamma}{\sigma_u}\right)}{\Phi\left(\frac{w_i\gamma}{\sigma_u}\right)}\right] = \ \rho\sigma_\varepsilon\lambda_i\,(\alpha_u)$$

con lo que $\rho\sigma_\varepsilon = \beta_*$ será un parámetro más a estimar en el modelo.

Ahora la ecuación resultado será:

$$Y_i = \ X_i\beta + \beta_*\lambda_i(\alpha_u) + \nu_i$$

Si en la estimación omitimos el segundo término los estimadores serían sesgados e inconsistentes, dependiendo el sesgo de la correlación ρ entre las perturbaciones ε_i, u_i .

Como aclaración recordaremos que el Ratio inverso de Mills es la relación entre la función de densidad y la función de distribución acumulada. Ocurre que si X es una variable que se distribuye como una normal con media μ y varianza σ^2

$$E\ (x/x\ > \alpha) = \mu + \sigma\frac{\phi\left(\frac{\alpha-\mu}{\sigma}\right)}{1-\Phi\left(\frac{\alpha-\mu}{\sigma}\right)}$$

$$E\ (x/x\ < \alpha) = \mu + \sigma\frac{-\phi\left(\frac{\alpha-\mu}{\sigma}\right)}{\Phi\left(\frac{\alpha-\mu}{\sigma}\right)}$$

Siendo ϕ la función estándar de densidad de una normal y Φ la distribución acumulada.

Heckman (1979) lo que propuso fue realizar en una primera etapa una regresión probit para los resultados positivos de la variable $Z_i = 1$, y utilizar los parámetros estimados para calcular el Ratio inverso de Mills, que se incluyen en la segunda etapa como variable explicativa de la ecuación de interés, estimándose ésta por MCO.

4. Descomposición de Oaxaca-Blinder, Neumark y Heckman

Si ahora retomamos la descomposición de Oaxaca-Blinder y aplicamos los resultados obtenidos en el epígrafe anterior, habrá que considerar que ante la existencia de sesgo de selección las dos ecuaciones de resultado deberán incluir el sesgo de selección, con lo que deberán reformularse como:

$$Y_A = X_A\beta_A + \lambda_A\beta_{\lambda A} + \nu_A$$

$$Y_B = X_B\beta_B + \lambda_B\beta_{\lambda B} + \nu_B$$

y la ecuación [1] será ahora:

$$R = (\bar{X}_A - \bar{X}_B)\hat{\beta}_A + \bar{X}_B\left(\hat{\beta}_A - \hat{\beta}_B\right) + \left(\hat{\lambda}_A\beta_{\lambda A} - \hat{\lambda}_B\beta_{\lambda B}\right)$$

o más intuitivamente, la diferencia entre medias se corrige por el sesgo de selección

$$(\bar{Y}_A - \bar{Y}_B) - \left(\hat{\lambda}_A\hat{\beta}_A - \hat{\lambda}_B\hat{\beta}_B\right) = (\bar{X}_A - \bar{X}_B)\hat{\beta}_A + \bar{X}_B\left(\hat{\beta}_A - \hat{\beta}_B\right)$$

o si incluimos el planteamiento de Neumark

$$(\bar{Y}_A - \bar{Y}_B) - \left(\hat{\lambda}_A\hat{\beta}_A - \hat{\lambda}_B\hat{\beta}_B\right) = \Delta\bar{X}\hat{\beta}_O + \left[\bar{X}_A\left(\hat{\beta}_A - \hat{\beta}_O\right) - \bar{X}_B\left(\hat{\beta}_B - \hat{\beta}_O\right)\right]$$

En un artículo posterior, Neumark y Oaxaca (2004) proponen utilizar:

$$\bar{Y}_A - \bar{Y}_B = R = (\bar{X}_A - \bar{X}_B)\,\hat{\beta}_A + \left(\hat{\lambda}_A - \hat{\lambda}_B^A\right)\hat{\beta}_{\lambda A} \text{ ……Diferencia características}$$

$$+\bar{X}_B\left(\hat{\beta}_A - \hat{\beta}_B\right) + \left(\hat{\lambda}_B^A - \hat{\lambda}_B\right)\hat{\beta}_{\lambda A} \text{ ....Discriminación}$$

$$+\hat{\lambda}_B\left(\hat{\beta}_{\lambda A} - \hat{\beta}_{\lambda B}\right)\text{..Selección}$$

y donde $\hat{\lambda}_B^A$ sería la inversa del ratio de Mills obtenida por las funciones de densidad y de distribución de los individuos del grupo B pero utilizando los parámetros del proceso de clasificación o selección del grupo A.

5. Descomposición de Oaxaca-Blinder en modelos Logit. Aproximación de Fairlie

De forma general, la ecuación o modelo resultado que se intenta descomponer en dos grupos no tiene porqué ser lineal. Este es el caso de utilizar un modelo Logit o Probit, donde la variable endógena es una variable binaria que es explicada por un conjunto de variables independientes mediante una relación logística.

Si recordamos que la descomposición original planteada en [1] se deducía de:

$$\bar{Y}_A - \bar{Y}_B = \bar{X}_A\,\hat{\beta}_A - \bar{X}_B\;\hat{\beta}_A + \bar{X}_B\;\hat{\beta}_A - \bar{X}_B\;\hat{\beta}_B$$

siguiendo el desarrollo propuesto por Fairlie (1999) tendremos que en el caso de que la relación $Y = F\left(X\hat{\beta}\right)$ sea no lineal, la anterior expresión tomará la forma:

$$\bar{Y}_A - \bar{Y}_B = \left[\sum_{i=1}^{N_A}\frac{F\left(X_{Ai}\hat{\beta}_A\right)}{N_A} - \sum_{i=1}^{N_B}\frac{F\left(X_{Bi}\hat{\beta}_A\right)}{N_B}\right] + \left[\sum_{i=1}^{N_B}\frac{F\left(X_{Bi}\hat{\beta}_A\right)}{N_B} - \sum_{i=1}^{N_B}\frac{F\left(X_{Bi}\hat{\beta}_B\right)}{N_B}\right] \quad [4]$$

evitando así el hecho de que al no ser lineal la relación pueda ocurrir que:

$$\bar{Y} \neq F\left(\bar{X}\hat{\beta}\right)$$

Para el caso de un modelo Logit o Probit, $\bar{Y}_j$ será la probabilidad media o porcentaje de unos para el grupo "j", y "F" será la función de distribución acumulada de la distribución Logit o de la distribución normal.

Al igual que en el modelo inicial de Oaxaca-Blinder este planteamiento tiene el mismo problema de seleccionar el grupo de referencia pues puede ser tanto el grupo *A* como el *B*. Por ello la ponderación realizada por Neumark es igualmente válida en este contexto.

Tal y como plantea Fairlie (2005) si admitimos que los dos grupos tienen tamaños iguales, $N_A = N_B$, se pueden utilizar todos los datos para estimar una regresión Logit de datos mezclados para obtener $\hat{\beta}$. Entonces es fácil calcular *F* y por ejemplo, en un modelo de dos variables la contribución de la variable X_1 a la diferencia entre los grupos será:

$$\hat{D}_1 = \frac{1}{N_B}\sum_{i=1}^{N_B} F\left(\hat{\alpha}_o + X_{A1i}\hat{\beta}_{1o} + X_{A2i}\hat{\beta}_{2o}\right) - F\left(\hat{\alpha}_o + X_{B1i}\hat{\beta}_{1o} + X_{A2i}\hat{\beta}_{2o}\right) \quad [5]$$

y la contribución de la variable X_2

$$\hat{D}_2 = \frac{1}{N_B}\sum_{i=1}^{N_B} F\left(\hat{\alpha}_o + X_{B1i}\hat{\beta}_{1o} + X_{A2i}\hat{\beta}_{2o}\right) - F\left(\hat{\alpha}_o + X_{B1i}\hat{\beta}_{1o} + X_{B2i}\hat{\beta}_{2o}\right) \quad [6]$$

donde:

X_{B1i} es la variable X_1 para el grupo *B*
X_{A1i} es la variable X_1 para el grupo *A*
X_{B2i} es la variable X_2 para el grupo *B*
X_{A2i} es la variable X_2 para el grupo *A*

El subíndice cero indica que los coeficientes han sido estimados para el total de la muestra y son un promedio de los dos grupos.[1]

Fairlie, siguiendo a Oaxaca y Ransom (1998), estima la varianza de la contribución de cada variable mediante:

[1] El orden de selección de las variables incide sobre la importancia de la contribución.

$$Var\left(\widehat{D}_1\right) = \left(\frac{\delta \widehat{D}_1}{\delta \hat{\beta}_o}\right)' Var\left(\widehat{B}_o\right)\left(\frac{\delta \widehat{D}_1}{\delta \hat{\beta}_o}\right)$$

donde [5] se ha reescrito por:

$$\widehat{D}_1 = \frac{1}{N_B}\sum_{i=1}^{N_B} F\left(X_{AAi}\hat{\beta}_o\right) - F\left(X_{BAi}\hat{\beta}_o\right)$$

y

$$\frac{\delta \widehat{D}_1}{\delta \hat{\beta}_O} = \frac{1}{N_B}\sum_{i=1}^{N_B} f\left(X_{AAi}\hat{\beta}_o\right)X_{AAi} - f\left(X_{BAi}\hat{\beta}_o\right)X_{BAi}$$

siendo f la función de densidad de la logística.

Dado que los tamaños muestrales raramente son iguales para los dos grupos Fairlie propone usar la muestra mezclada o total para calcular las estimaciones $\widehat{Y}_i$, tanto para el grupo A como para el B. Si el grupo B es más pequeño que el grupo A, se tomará una submuestra del A de igual tamaño que el B y se realizará la estimación y descomposición para ambos grupos. El procedimiento se puede repetir con múltiples submuestras y de aquí obtener los valores promedio de la descomposición. Fairlie tomó un total de 1.000 submuestras.

En su primer trabajo Fairlie (1999) aplica el modelo a las diferencias de autoempleo entre blancos y negros en Estados Unidos (Tabla 1). Aquí se recogen los resultados de la modelización de la probabilidad de convertirse en autoempleado, si bien también estima otro modelo para la probabilidad de abandono del autoempleo. En este caso la diferencia de probabilidad entre negros y blancos es un 2% y esta diferencia se explica en un 29,1% por las variables incluidas cuando se toman los negros como grupo de referencia y del 14,2% cuando se toman los blancos. Si bien no presenta los resultados, estima la regresión con los datos mezcla de blancos y negros en la forma planteada por Neumark, encontrándolos muy próximos a la segunda especificación en la que el grupo de referencia son los blancos.

Tabla 1

Decomposition of Difference between White and Black Transition Rates into Self-Employment: PSID (1968–1989)

	Specification	
	(1)	(2)
Sample used to estimate coefficients and transition rates	Black	White
Transition rate (gap = .0201)	.0193	.0394
Contribution to the gap from racial differences in the following variables:		
1. Controls	−.0010	−.0034
	−5.0%	−17.1%
2. Education level	.0001	.0012
	.5%	5.8%
3. Asset levels	.0031	.0028
	15.2%	13.9%
4. Father's education level	.0010	.0008
	4.8%	4.2%
5. Father's self-employment status	.0027	.0015
	13.6%	7.5%
6. All variables	.0058	.0029
	29.1%	14.2%

NOTE.—The sample consists of male nonagricultural workers (ages 16–65) who are heads of family units. Contribution estimates are mean values of the decomposition using 1,000 random subsamples of whites. Sample weights provided by the PSID are used in all calculations. Controls include all variables included in the logit regressions but not listed separately in this table.

En un posterior trabajo, Fairlie (2005), mantiene su planteamiento y realiza una serie de análisis y comparaciones con otros procedimientos. En primer lugar compara su estimación no lineal con la resultante de aplicar la descomposición de Oaxaca-Blinder a un modelo lineal de probabilidad, encontrando que para algunas aplicaciones en las que el efecto es importante, Oaxaca-Blinder tiende a aumentar la contribución fuera de rangos aceptables. Por otro lado reconoce que el ordenamiento de las variables en el modelo puede afectar a los resultados, por lo que propone experimentar con la ordenación y determinar si ésta afecta en una aplicación concreta. Si se apreciaran cambios importantes en la contribución de las variables la ordenación puede ser aleatoria y usando un número elevado de simulaciones encontrar la contribución media de cada variable. (Ver tabla 2)

Tabla 2

Non-Linear Decompositions of Black/White Gaps in Home Computer Rates
Using Various Coefficient Estimates

	Specification			
	(1)	(2)	(3)	(4)
Sample used for coefficients	White	Black	Black/White Pooled	All Races Pooled
White computer ownership rate	0.7286	0.7286	0.7286	0.7286
Black computer ownership rate	0.4257	0.4257	0.4257	0.4257
Black/White gap	0.3030	0.3030	0.3030	0.3030
Contributions from racial differences in:				
Sex and age	-0.0004	0.0001	-0.0004	-0.0002
	(0.0003)	(0.0010)	(0.0003)	(0.0002)
	-0.1%	0.0%	-0.1%	-0.1%
Marital status and children	0.0315	0.0302	0.0317	0.0289
	(0.0017)	(0.0041)	(0.0015)	(0.0014)
	10.4%	10.0%	10.5%	9.5%
Education	0.0340	0.0367	0.0341	0.0369
	(0.0010)	(0.0028)	(0.0009)	(0.0008)
	11.2%	12.1%	11.2%	12.2%
Income	0.0775	0.1031	0.0797	0.0799
	(0.0020)	(0.0048)	(0.0019)	(0.0017)
	25.6%	34.0%	26.3%	26.4%
Region	0.0102	0.0157	0.0104	0.0085
	(0.0017)	(0.0055)	(0.0016)	(0.0015)
	3.4%	5.2%	3.4%	2.8%
Central city status	-0.0015	0.0000	-0.0007	-0.0019
	(0.0021)	(0.0041)	(0.0019)	(0.0017)
	-0.5%	0.0%	-0.2%	-0.6%
All included variables	0.1512	0.1859	0.1548	0.1519
	49.9%	61.3%	51.1%	50.2%

Notes: (1) The sample consists of adults ages 25-55 from the specified racial group/s. (2) Standard errors are reported in parantheses below contribution estimates. (3) The sample sizes used to estimate the coefficients in Specifications 1-4 are 34,386, 4,555, 38,941 and 46,322, respectively. (4) Contribution estimates are mean values of the decomposition using 1000 subsamples of whites. See text for more details.

6. Método general de Yun para modelos no lineales (2004)

Yun (2004) propone un método alternativo de descomposición para modelos no lineales al que denomina "Método Generalizado de Descomposición de Diferencias en el Primer Momento", es decir con respecto a la media. Al estar trabajando con un modelo no lineal, pongamos que probit, el modelo tiene la expresión general:

$$Prob(Y)_g \; = \; \Phi\left(X_g B_g\right)$$

siendo Φ la función de distribución acumulada de la normal estándar, g el subíndice que representa a los dos grupos, X el vector de variables explicativas y B el vector de parámetros. La descomposición de Oaxaca-Blinder podemos plantearla para valores medios por:

$$\overline{Prob(Y)_A} - \overline{Prob(Y)_B} = \left[\overline{\Phi(X_B\beta_B)} - \overline{\Phi(X_A\beta_B)}\right] + \left[\overline{\Phi(X_A\beta_B)} - \overline{\Phi(X_A\beta_A)}\right]$$

Con lo que se tratará de calcular los valores medios de las funciones de distribución. El análisis requiere determinar la contribución de cada variable a las diferencias de las probabilidades de los dos grupos. Para obtener tal contribución Yun propone una transformación en dos etapas. En la primera se estima la función para los valores medios de los regresores y en la segunda se efectúa una expansión de Taylor de primer orden en la que se obtienen los valores y errores de la función sobre $\bar{X}_A\beta_A$ y $\bar{X}_B\beta_B$. La expresión final propuesta por Yun es:

$$\overline{Prob(Y)_A} - \overline{Prob(Y)_B} = \sum_{i=1}^{k} P_{\Delta X}^{i}\left[\overline{\Phi(X_B\beta_B)} - \overline{\Phi(X_A\beta_B)}\right] + \sum_{i=1}^{k} P_{\Delta X}^{i}\left[\overline{\Phi(X_A\beta_B)} - \overline{\Phi(X_A\beta_A)}\right]$$

con

$$P_{\Delta X}^{i} = \frac{\left(\bar{X}_B^{i} - \bar{X}_A^{i}\right)\beta_B^{i}}{(\bar{X}_B - \bar{X}_A)\beta_B}$$

$$P_{\Delta B}^{i} = \frac{\bar{X}_A^{i}\left(\beta_B^{i} - \beta_A^{i}\right)}{\bar{X}_A(\beta_B - \beta_A)}$$

$$\sum_{i=1}^{k} P_{\Delta X}^{i} = \sum_{i=1}^{k} P_{\Delta \beta}^{i} = 1$$

De esta forma los P^i recogen la influencia o peso de las variables y coeficientes sobre la descomposición inicial y el subíndice $i = 1, 2 \dots \dots k$, se refiere a las variables incluidas en el modelo.

Con este método es posible calcular la descomposición inicial, estableciendo la parte correspondiente a la diferencia en variables y la parte correspondiente a los parámetros. Posteriormente se calcula la importancia relativa de cada variable y de cada parámetro en cada una de las diferencias de la descomposición.

7. Utilización de Stata en modelos no lineales

Tanto la aproximación de Fairlie como la de Yun tienen por objetivo estimar la descomposición de Oaxaca-Bilder cuando las relaciones son no lineales, siendo un caso especial cuando la variable endógena es una variable binaria. Dado que el caso de la no linealidad ha sido más recientemente introducido y no aparece en los manuales como el de Cameron y Trivedi (2009), describiremos a continuación para el caso de dos variables explicativas y una variable clasificatoria del grupo, los comandos necesarios para su estimación con el programa Stata. En ambos casos se realiza una estimación Logit mediante:

Para el caso de Fairlie:

```
fairlie Y X1 X2, by (GRUPO)
```

Para el caso de Yun:

```
oaxaca  Y X1 X2, by (GRUPO) logit pooled
```

dónde Y es una variable binaria y X1 y X2 son dos variables explicativas que pueden ser tanto binarias como cuantitativas.

Tal y como se expuso en los anteriores apartados la utilización de la metodología de Yun obtiene la contribución de parámetros y variables a la diferencia entre los grupos, mientras que la de Fairlie obtiene la contribución de las variables.

BIBLIOGRAFÍA

- **Albrecht, James W., Bjorklund, Anders and Vroman, Susan B. (2001)**, “Is There a Glass Ceiling in Sweden” (April 2001). IZA Discussion Paper No. 282.

- **Albrecht, James W., Nguyen, Binh and Vroman, Susan B. (2003),** "A Quantile Regression Decomposition of Urban-Rural Inequality in Vietnam," Working Papers gueconwpa-03-03-31, Georgetown University, Department of Economics.

- **Blinder, Alan S. (1973),** “Wage Discrimination: Reduced Form and Structural Estimates” The Journal of Human Resources, Vol. 8, No. 4, pp. 436-455 Autumn, 1973.

- **Cameron, A.L., Triverdi, P.K. (2009)**, “Microeconometrics using Stata”. Stata Corp LP.

- **Cotton Jeremiah (1998),** ‘‘On the Decomposition of Wage Differentials,’’ Review of Economics and Statistics 70 (1988), pp. 236–243.

- **Elder, Todd E., Goddeeris, John H. and Haider, Steven J., (2010)**, "Unexplained Gaps and Oaxaca-Blinder Decompositions," Labour Economics, Vol. 17 (1) 2010.

- **Fairlie, R. (1999), “**The absence of African-American owned business: an analysis of the dynamics of selfemployment” Journal of Labor Economics, 17, pp. 80–108.

- **Fairlie, R. (2005), “**An extension of the Blinder-Oaxaca decomposition technique to logit and probit models”. Journal of Economic and Social Measurement, Vol 30, pp. 305–316.

- **Greene, W. H. (2003)**, “Econometric Analysis”. Fifth Edition. Prentice Hall

- **Heckman, J. (1979),** "Sample Selection Bias as a Specification Error," Econometrica, Econometric Society, vol. 47(1), pp. 153-61, January.

- **Mahadevan, R. y Suardi, S. (2012),** "Impact of socio-economic factors and social affiliation on living standards: a quantile regression aproach". Applied Economics Letters, 2012, Vol 19, pp. 1.231-1.236.

- **Oaxaca, R. L. (1973),** "Male-female wage differentials in urban labor markets," International Economic Review, Vol 14(3), pp. 693-709.

- **Oaxaca, R. L., and Michael R. Ransom (1988),** ''Searching for the Effect of Unionism on the Wages of Union and Non-Union" Workers, Journal of Labor Research 9 (1988), pp. 139–148.

- **Oaxaca, R. L., and Michael R. Ransom (1994),** "On Discrimination and the Decomposition of Wage Differentials,'' in Shoshana Neuman and Jacques Silber (eds), Journal of Econometrics, Vol. 61(1), pp. 5-21, March.

- **Oaxaca, R., and Michael Ransom (1998),** "Calculation of Approximate Variances for Wage Decomposition Differentials," Journal of Economic and Social Measurement, 24, pp. 55-61.

- **Neumark, D. (1988), "**Employers discriminatory behavior and the estimation of wage discrimination" (The Journal of Human Resources, 23, pp. 279–295)

- **Reimers, C. (1983),** ''Labor Market Discrimination Against Hispanic and Black Men', The Review of Economics and Statistics, Vol 65 nº 4 (1983), pp.570–579.

- **Yun, Myeong-Su, (2004**), "Decomposing differences in the first moment," Economics Letters, Elsevier, vol. 82(2), pages 275-280, February.