## Capítulo 2
## Marco teórico.

Además de tratarse de una cuestión que por si misma representa efectos negativos para una sociedad,[1] la desigualdad juega un papel preponderante en el desarrollo económico de los países. Bourguignon (2002) concluye que con una redistribución de ingresos se reduce la pobreza a través de un incremento en la elasticidad de la reducción de pobreza con respecto al crecimiento.[2] Además, si se combina, por ejemplo, la desigualdad entre ingresos con la imperfección en los mercados financieros, se limita la capacidad de los pobres para adquirir bienes tales como capital humano, tierra, y vivienda, lo que reduce sus oportunidades futuras y la posibilidad de suavizar consumo ante la presencia de cualquier choque en la economía, lo que reduce el crecimiento y el desarrollo en general.[3]

Por ello es importante contar con información sobre los factores que intervienen en la dinámica de la distribución del ingreso en un país para una mejor comprensión de su proceso de desarrollo e implementación de política económica en materia distributiva.[4] Sin embargo, las técnicas estadísticas de descomposición para obtener esta información han estado limitadas a estudiar diferencias en medidas resumen de distribuciones en vez de analizar distribuciones totales, enfocándose en definiciones específicas de bienestar social en lugar de distribuciones de bienestar individual. Los ejemplos más representativos son la descomposición Oaxaca-Blinder de las diferencias entre los ingresos medios de grupos poblacionales con características diferentes y las medidas

[1] De acuerdo a De Ferranti et al. (2003), la aversión a la desigualdad, o preferencia por la equidad, ha sido un punto de vista dominante (aunque no consensuado) dentro de la filosofía política y de las teorías de justicia social. Estos autores también mencionan aspectos de economía política que consisten en que en las sociedades con alta concentración de poder y riqueza, las elites tienen mayor libertad de acción en elegir estrategias que beneficien a ellos mismos en lugar de los grupos medios o bajos. Las sociedades desiguales en las cuales el poder político se entrecruza con la riqueza, elegirán con menor probabilidad políticas que reduzcan esas ineficiencias y asignarán recursos a usos alternativos, lo cual puede conducir a una disminución en el crecimiento.

[2] Kanbur y Lustig (1999) mencionan que, a diferencia de la década de los ochenta, a partir de los noventa el estudio de la desigualdad cobró gran importancia debido a, entre otras cosas, su relación con los niveles de desarrollo de los países.

[3] De Ferranti et al. (2003). Para Kanbur y Lustig (1999) otra razón para que la desigualdad retomara importancia fue que “las últimas dos décadas culminaron con un triunfo de la literatura que estudia el papel de los mercados y de la información imperfectos”.

[4] Las páginas siguientes se basan en Bourguignon et al. (2004).

resumen de desigualdad con propiedades de descomposición. En ambos casos la lógica es que el ingreso medio o desigualdad en una población son resultado de la agregación de varios grupos socio-demográficos o fuentes de ingresos, con lo que es posible explicar los cambios en la totalidad de los ingresos medios o en las medidas de desigualdad identificando cambios en las medias y desigualdad dentro de tales grupos o fuentes de ingresos. Así, al descomponer medidas de desigualdad como el coeficiente de Theil o la desviación media de los logaritmos del ingreso de acuerdo a género, educación o grupos de edad es posible obtener información acerca de cambios estructurales en la distribución de los ingresos de un país.

No obstante, existe la necesidad de trabajar basándonos en distribuciones totales en vez de primeros momentos o índices de desigualdad. En particular, los estudios sobre pobreza y aspectos de desarrollo requieren la capacidad de analizar cómo el cambio en la distribución de los ingresos puede cambiar la cercanía de un país con respecto a su línea de pobreza. Por ejemplo, en términos del enfoque Oaxaca-Blinder la cuestión no sería tanto saber si los ingresos de las mujeres son menores que los de los hombres debido a que aquéllas tienen menores niveles de educación, sino si es más o menos así en la parte baja de la distribución de ingresos. Responder este tipo de preguntas requiere trabajar con distribuciones totales en lugar de medidas resumen. Varias técnicas para descomponer cambios distributivos en lugar de cambios en las medidas de desigualdad han sido desarrolladas desde la década pasada o más, en parte debido a mayores avances computacionales. La metodología utilizada en el presente trabajo para analizar cambios distributivos tiene su marco teórico en dichas técnicas de descomposición, las cuales se exponen a continuación.

### 2.1. Métodos escalares de descomposición de cambios distributivos.

#### 2.1.1. Descomposición de cambios en medias de los ingresos: el Método Oaxaca-Blinder.

Oaxaca (1973) y Blinder (1973) establecieron la siguiente forma de comparar los ingresos medios de dos poblaciones distintas.[5] Supóngase que el ingreso se estima con el siguiente modelo lineal en los periodos *t* y *t'*:

$$
\begin{aligned}
y_{it} &= \beta_t X_{it} + u_{it} \\
y_{jt'} &= \beta_{t'} X_{jt'} + u_{jt'}
\end{aligned}
\qquad (1)
$$

Este modelo supone que los ingresos de los individuos *i* en el periodo *t* dependen linealmente de un vector con sus características observadas, $X_{it}$, y de algunas características no observadas capturadas en el término residual, $u_{it}$. La misma relación se mantiene para los individuos *j* en el periodo *t'*, quienes presumiblemente son distintos a aquéllos observados en el periodo *t*. Los coeficientes $\beta_t$ y $\beta_{t'}$ indican el efecto de las características individuales, *X*, en los ingresos, *y*. Si los componentes de *X* son interpretados como 'dotaciones' o 'activos' individuales, entonces las $\beta$'s pueden tomarse como tasas de retorno o los 'precios' que paga el mercado de acuerdo a dichos activos.

Dadas dos muestras de individuos, una en el periodo *t* y otra en el periodo *t'*, estos retornos pueden estimarse por Mínimos Cuadrados Ordinarios, bajo el supuesto común de que los términos residuales son independientes de las dotaciones observadas.

Considérese ahora el cambio en los ingresos medios del periodo *t* al *t'*. Bajo el supuesto inocuo de que el valor esperado de los residuales es cero, mediante una transformación se llega a la siguiente descomposición del cambio en las medias:

[5] Estos autores estaban interesados en la discriminación de salarios entre individuos de acuerdo a características como género o raza. Conceptualmente, no hay diferencia entre considerar dos poblaciones en dos puntos en el tiempo, como se hace más adelante.

$$\Delta\bar{y} = \bar{y}_{t'} - \bar{y}_t = \beta_t(\bar{X}_{t'} - \bar{X}_t) + \bar{X}_{t'}(\beta_{t'} - \beta_t) \qquad (2)$$

El cambio en los ingresos medios aparece como la suma de dos efectos: a) el del cambio en las dotaciones medias a precios constantes (‘efecto activos’), y b) el cambio en los retornos con las dotaciones medias del periodo *t’* (‘efecto precios’). En otras palabras, el cambio en el ingreso medio de la población entre los periodos *t* y *t’* es explicado por un cambio en sus características (educación, edad, área de residencia, etc.) y por un cambio en las tasas de retorno de tales características. Por ejemplo, cuando la descomposición Oaxaca-Blinder analiza las diferencias de género, el diferencial hombre/mujer de ingresos medios se descompone en lo que se debe a que: a) las mujeres y hombres que trabajan no tienen las mismas características en términos de educación, edad u ocupación, y b) a las mujeres y a los hombres no se les paga la misma tasa de retorno, manteniendo las características constantes.

El interés práctico de una descomposición como la de la ecuación (2) radica en que si el análisis económico es capaz de predecir o explicar cambios en el sistema de precios, $\beta$, entonces será posible describir las implicaciones de estos cambios en la evolución de los ingresos medios individuales. Debemos tomar en cuenta que esta descomposición ignora cualquier relación entre las dos fuentes de cambio. Es decir, es probable que los cambios observados en los retornos se deban a cambios en la estructura socio-demográfica de la población, y viceversa. Por ejemplo, una fuerza de trabajo más educada puede llevar a brechas en los retornos menores, y una brecha en los retornos alta puede ser un incentivo para que parte de la población se vuelva más educada.

Debe mencionarse un aspecto adicional sobre la descomposición Oaxaca-Blinder, el cual tiene que ver con que la ecuación (2) de descomposición depende del periodo base que se tome[6]. Es decir, una identidad similar a (2) es:

$$\Delta\bar{y} = \bar{y}_{t'} - \bar{y}_t = \beta_{t'}(\bar{X}_{t'} - \bar{X}_t) + \bar{X}_t(\beta_{t'} - \beta_t) \qquad (3)$$

---

[6] En la literatura esto se conoce como *path dependent*, debido a que se refiere al papel de la ‘senda’ o ‘dirección’ que toma la descomposición.

En este caso, el sentido o la dirección de la descomposición indica que el 'efecto activos' es evaluado utilizando los precios del periodo *t'*, mientras que el 'efecto precios' se estima usando las dotaciones de activos medias iniciales. No hay razón para que se obtengan los mismos estimadores con esta descomposición de los efectos precios y activos que los que se obtuvieron con (2). De acuerdo a esto, la dirección o senda utilizada para la descomposición debe tomarse en cuenta en el análisis.[7]

### 2.1.2. Descomposición de medidas de desigualdad de ingresos.

La descomposición del apartado anterior puede aplicarse también a medidas resumen de desigualdad de Entropía Generalizada conocidas como 'descomponibles',[8] dado que tienen propiedades de descomposición muy deseables.[9]

Supóngase que la población de individuos perceptores de ingreso está dividida en $G$ grupos, $g = 1, 2, ..., G$, y que la medida de desigualdad en el grupo $g$ está denotada por $I_g$, donde $I$ es la desigualdad del total de la población. Tales medidas satisfacen la siguiente propiedad general:

$$I = \sum_{g=1}^{G} I_g w(n_g, m_g) + \bar{I}(n_1, \bar{y}_1; n_2, \bar{y}_2; ...; n_G, \bar{y}_G) = I_W + I_B \qquad (4)$$

donde $n_g$ y $m_g$ corresponden respectivamente a las proporciones de población e ingreso del grupo $g$ dentro de la población total e $\bar{I}$ ( ) es la desigualdad existente *entre* los grupos $g$. La distribución del ingreso consiste en $n_1$ veces el ingreso $\bar{y}_1$, $n_2$ veces el ingreso $\bar{y}_2$, etc. La desigualdad total, $I$, se descompone entonces en dos términos: la media *dentro* de la desigualdad de grupos, donde cada grupo $g$ es ponderado por un peso, $w$, el cual depende de las proporciones de población e ingreso, y la desigualdad *intra* grupos, $\bar{I}$ ( ).

[7] Más adelante se explica la forma en que se tomó en cuenta este aspecto dentro de la aplicación empírica.
[8] Traducción del término en inglés 'descomposable'.
[9] Para una introducción de medidas de desigualdad descomponibles, ver por ejemplo Cowell (2002) y las referencias que ahí se exponen.

Sin embargo, no es la descomposición entre grupos en un punto en el tiempo lo que nos interesa sino la descomposición del cambio en la desigualdad entre dos puntos en el tiempo. Diferenciando (4), el cambio en la desigualdad total, $\Delta I$, podría expresarse como la suma del cambio en la desigualdad dentro de los grupos, $\Delta I_W$, y el cambio en la desigualdad intra grupos, $\Delta I_B$. Para ello podemos tomar la *desviación del logaritmo*, que es la más simple de las medidas descomponibles de desigualdad. Dada una población de *n* individuos *i*, la desviación del logaritmo es:

$$L = \sum_{i=1}^{n} \frac{1}{n} Log(\bar{y} / y_i) \qquad (5)$$

En este caso, la ecuación (4) se escribe:

$$L = \sum_{g=1}^{G} n_g L_g + \sum_{g=1}^{G} n_g Log(\bar{y}_g / \bar{y}) = I_W + I_B \qquad (6)$$

Finalmente, diferenciando esta expresión entre dos periodos, *t* y *t'*, se tiene:

$$\Delta L \approx \sum_{g=1}^{G} n_g \left[ \frac{\Delta \bar{y}}{\bar{y}} - \frac{\Delta \bar{y}_g}{\bar{y}_g} \right] + \sum_{g=1}^{G} \left[ L_g + Log(\bar{y} / \bar{y}_g) \right] \Delta n_g + \sum_{g=1}^{G} n_g \Delta L_g \qquad (7)$$

El cambio total en la desigualdad es expresado entonces como la suma de tres tipos de efectos: a) el cambio en el ingreso medio relativo del grupo, b) el cambio en las ponderaciones poblacionales de los grupos, y c) el cambio en la desigualdad intra grupos. Expresiones análogas pueden derivarse para los otros miembros de la familia de medidas de desigualdad descomponibles, como los índices de Gini y de Theil.

Para ilustrar lo anterior, supóngase una población de perceptores de ingreso que ha sido dividida en grupos de acuerdo a la educación: individuos sin escolaridad, con primaria, con secundaria incompleta, etc. El cambio en la desigualdad total entre los años *t* y *t'* es la suma de: a) los efectos de cambios en los ingresos relativos por nivel educativo, b) los efectos de cambios en la estructura educativa de la población, y c) cambios en la desigualdad intra grupos educativos. Este último término es comúnmente calculado como un *residual*, y corresponde a aquella parte del cambio en la desigualdad que no es explicada por el cambio en los ingresos medios entre grupos educativos ni por la estructura educativa de la población.

La descomposición anterior puede implementarse para todas las características posibles de la población, así como para todas las combinaciones posibles de tales características. Por ejemplo, los grupos pueden ser definidos simultáneamente de acuerdo a la educación del jefe del hogar, las edades, la zona de residencia (rural o urbana), el número miembros del hogar, etc.[10] Una de las razones del atractivo de esta metodología es su analogía con la descomposición Oaxaca-Blinder: los cambios en los ingresos medios relativos de los grupos juegan un papel similar al cambio en los coeficientes de los 'precios', $\beta$, mientras que el cambio en las ponderaciones de los grupos es otra forma de representar los cambios en la estructura socio-demográfica de la población, $\overline{X}_{t'} - \overline{X}_t$.

Hay dos diferencias entre estos dos enfoques, más allá del hecho de que uno se refiere al ingreso medio y el otro a la desigualdad entre ingresos. Primero, la fórmula de descomposición de la desigualdad es no-paramétrica, mientras que la de Oaxaca-Blinder se desarrolla en un modelo de ingreso lineal.[11] Segundo, la descomposición de la desigualdad tiene un término residual correspondiente al cambio en la desigualdad intra grupos, el cual, a pesar del hecho de que no permite por si mismo una fácil interpretación económica, en aplicaciones empíricas se vuelve un componente importante del cambio observado en la desigualdad.

Una limitante de la descomposición descrita en este apartado es que se aplica a la desagregación de la población en subgrupos pero no en la desagregación de ingresos por

[10] Bourguignon et al. (2004) mencionan que se han realizado numerosas aplicaciones de esta metodología de descomposición, comenzando con el análisis de la evolución de la desigualdad en el Reino Unido de Mookherjee y Shorrocks (1983).
[11] Aunque el método Oaxaca-Blinder puede también llevarse a cabo en términos de las medias de los grupos y pesos, y no en términos de un modelo de ingreso lineal.

fuentes. Supóngase que el ingreso del individuo *i* puede expresarse como la suma de los ingresos provenientes de las fuentes *1* –por ejemplo salarios– y *2* –ingreso por autoemplearse– :

$$y_i = y_{1i} + y_{2i} \quad (8)$$

Podría ser de interés descomponer el cambio de la desigualdad en el ingreso total en lo que fue debido al cambio en la media y en las fuentes de desigualdad de los ingresos *1* y *2*. La fórmula de descomposición anterior no funciona en ese caso. En particular, simplemente no es cierto que la desigualdad total sea el promedio ponderado de la desigualdad en cada fuente de ingreso. Es importante entonces la manera en que las dos fuentes *covarían* dentro de la población.

Shorrocks (1982) muestra que la desigualdad total $I_y$ en un punto en el tiempo puede ser descompuesta en la desigualdad proveniente de varias fuentes de ingreso. Este autor demuestra que para cualquier medida de desigualdad, es igualmente cierto que:

$$I_y = \frac{\operatorname{cov}(y_1, y)}{V(y)} I_y + \frac{\operatorname{cov}(y_2, y)}{V(y)} I_y \quad (9)$$

Donde $\operatorname{cov}(y_j, y)$ es la covarianza entre la fuente de ingreso ($j = 1,2$) y el ingreso total en la población. En otras palabras, la razón de esta covarianza y la varianza del ingreso total puede ser interpretada como el porcentaje de contribución de la fuente de ingreso *j* en la desigualdad total, independientemente de la medida de desigualdad que se esté usando.

Esta descomposición se complica cuando son considerados cambios en el tiempo. De hecho, con el objetivo de analizar cómo un cambio en la distribución de una fuente de ingreso –dígase fuente *1*– puede modificar la desigualdad en el ingreso total, primero es necesario calcular cómo este cambio puede modificar la covarianza entre esa fuente de ingreso y el ingreso total. Esto requiere calcular cómo un cambio en la distribución de la fuente *1* puede por si misma modificar la covarianza entre los ingresos de las fuentes *1* y

*2*. Para este análisis debe conocerse no sólo el cambio en la distribución marginal del ingreso de una fuente, sino también el cambio en la distribución conjunta del ingreso proveniente de varias fuentes. La necesidad de utilizar esta distribución conjunta puede explicar porqué la propiedad anterior de descomposición por fuente de ingreso es muy poco usada en trabajos empíricos de cambios distributivos.

## 2.2. Descomposición de cambios distributivos: métodos paramétricos y no-paramétricos.

### 2.2.1. Los contrafactuales distributivos.

El análisis de descomposición escalar anterior puede extenderse al caso de cambios distributivos. Siendo $f^{t}(y)$ y $f^{t'}(y)$ las funciones de densidad de la distribución del ingreso, $y$, o cualquier otra definición de bienestar económico en los periodos $t$ y $t'$, el objetivo del análisis es identificar los factores que inciden en el cambio ocurrido entre la primera y la segunda distribución.

Para ello es posible partir de las distribuciones conjuntas $\varphi^{\tau}(y, X)$, donde $X$ es un vector de características observadas en individuos o familias, tales como edad, educación, ocupación, tamaño de la familia, etc. El superíndice $\tau$ (= $t$, $t'$) indica el periodo en que esta distribución conjunta es observada. La distribución de los ingresos de los hogares, $f^{\tau}(y)$ es entonces la distribución marginal de la distribución conjunta $\varphi^{\tau}(y,X)$:

$$f^{\tau}(y) = \int \ldots \int_{C(X)} \varphi^{\tau}(y, X)\, dX \qquad (10)$$

Estableciendo $g^{\tau}(y|X)$ como la distribución de ingreso condicional en $X$, una expresión equivalente de la distribución marginal del ingreso en el periodo $\tau$ es:

$$f^{\tau}(y) = \int \ldots \int_{C(X)} g^{\tau}(y \mid X)\, \chi^{\tau}(X)\, dX \qquad (11)$$

donde $\chi^{\tau}(X)$ es la distribución conjunta de todos los elementos de $X$ en el periodo $\tau$.

Dada esta descomposición es posible expresar el cambio distributivo observado de $f^{t}(\,)$ a $f^{t'}(\,)$ como función del cambio en las dos distribuciones que aparecen en la ecuación (11) (la del ingreso condicional a las características *X*, $g(y \mid X)$, y la distribución de tales características, $\chi(X)$). Para hacer eso, se define el siguiente experimento contrafactual:

$$f_g^{t \to t'}(y) = \int \ldots \int_{C(X)} g^{t'}(y \mid X)\chi^{t}(X)dX \qquad (12)$$

Esta es la distribución que hubiera sido observada en el periodo *t* si la distribución de ingresos condicionada a las características *X* hubiera sido la observada en el periodo *t'*. Esta distribución contrafactual puede ser fácilmente calculada una vez que las distribuciones condicionales $g^{t}(y|X)$ y $g^{t'}(y|X)$, así como la distribución marginal $\chi^{t}(X)$ han sido definidas. Puede definirse entonces el contrafactual:

$$f_\chi^{t \to t'}(y) = \int \ldots \int_{C(X)} g^{t}(y \mid X)\chi^{t'}(X)dX \qquad (13)$$

En esta ecuación fue la distribución conjunta de características la que se modificó. Nótese que esta última distribución pudo también haberse obtenido comenzando a partir del periodo *t'* y reemplazando la distribución del ingreso condicional de ese periodo por la observada en el periodo *t*. Esto puede expresarse:

$$f_g^{t \to t'}(y) \equiv f_\chi^{t' \to t}(y) \quad \text{y} \quad f_\chi^{t \to t'}(y) \equiv f_g^{t' \to t}(y) \qquad (14)$$

Basándonos en la definición de estos contrafactuales, el cambio distributivo observado $f^{t'}(y)$-$f^{t}(y)$ puede ahora descomponerse en:

$$f^{t'}(y) - f^{t}(y) \equiv \left[f_g^{t \to t'}(y) - f^{t}(y)\right] + \left[f^{t'}(y) - f_g^{t \to t'}(y)\right] \qquad (15)$$

Como en el caso del método Oaxaca-Blinder, el cambio distributivo observado es expresado como la suma de un efecto ‘precios’ y un efecto ‘activos’. El primer término del lado derecho de (15) describe la manera en la que la distribución del ingreso ha cambiado en el tiempo debido a un cambio en la distribución condicional de las características *X*. En otras palabras, muestra cómo la misma distribución de las características – las del periodo *t* – resulta en diferentes distribuciones del ingreso cuando se le compara en dos periodos. Para comprobar que el segundo término es el efecto de un cambio en la distribución de las características que tuvo lugar en los periodos *t* y *t’*, tomamos la ecuación (14) y reescribimos la fórmula anterior de descomposición como:

$$f^{t'}(y) - f^{t}(y) = \left[f_g^{t\to t'}(y) - f^{t}(y)\right] + \left[f^{t'}(y) - f_\chi^{t'\to t}(y)\right] \qquad (16)$$

La principal diferencia de este método con el de Oaxaca-Blinder y la descomposición de medidas escalares de desigualdad es que la fórmula de descomposición y los contrafactuales de los cuales depende se refieren a distribuciones totales en lugar de sus medias. Tomando las medias de (15) y (16) bajo el supuesto paramétrico de que la media condicional de $g^t(y|X)$ puede ser expresada como $\beta_\tau X$ se podría llegar a la ecuación (2) de Oaxaca-Blinder.

La única limitante en la descomposición anterior es la dependencia en la dirección o senda que tome discutida anteriormente en relación a la ecuación Oaxaca-Blinder, lo cual significa que al cambiar la distribución del ingreso condicional de la observada en *t* a la observada en *t’* no se obtiene el mismo efecto al cambiar la distribución en un sentido inverso. Eso quiere decir que:

$$\left[f_g^{t\to t'}(y) - f^{t}(y)\right] \neq \left[f^{t'}(y) - f_g^{t'\to t}(y)\right] \qquad (17)$$

Sin embargo, es probable que la diferencia sea pequeña cuando el cambio en las distribuciones de ingreso condicionales $g(y|X)$ es pequeño.

### 2.2.2. Aplicación de contrafactuales.

Resulta de mayor interés en algunos casos descomponer aún más el cambio en la distribución de las características individuales (*X*). Es decir, sería deseable separar, por ejemplo, el efecto del cambio en la distribución de la escolaridad o del tamaño de la familia. Esto requiere entonces extender la ecuación (11) para incorporarle un mayor número de condicionantes y definir nuevos contrafactuales.

Si clasificamos las variables *X* en (*V*,*W*), la cadena de condicionantes de (11) puede reescribirse:

$$f^{\tau}(y) = \int \ldots \int_{C(V,W)} g^{\tau}(y \mid V, W) h^{\tau}(V \mid W) \psi^{\tau}(W)\, dV dW \qquad (18)$$

donde $h^{\tau}(V \mid W)$ es la distribución de *V* condicional en *W* y $\psi^{\tau}(W)$ la distribución marginal de *W*. El conjunto de contrafactuales puede entonces ampliarse modificando la distribución condicional de *V*. Todas las combinaciones de las tres distribuciones, $g^{\tau_g}(y \mid V, W)$, $h^{\tau_h}(V \mid W)$ y $\psi^{\tau_\psi}(W)$ con $\tau_g, \tau_h, \tau_\psi = t$ o $t'$ pueden ser consideradas como generadoras de un contrafactual específico.

Comparando dos contrafactuales que difieren en sólo una distribución se obtiene un estimador de la contribución del cambio en esa distribución en particular en el cambio distributivo total. Obviamente, hay muchas direcciones (‘paths’) para evaluar esta contribución, sin la garantía de que lleven al mismo estimador.

Es posible obtener una descomposición condicionada más detallada de variables en *V*. Por ejemplo, podría ser de interés analizar el efecto distributivo de un cambio en la distribución de algunos componentes de *V* condicionales a los otros, con lo que habría que descomponer $h^{t}(V \mid W)$ en $h_1^{\tau}(V_1 \mid V_{-1}, W) h_{-1}^{\tau}(V_{-1} \mid W)$, donde $V_{-1}$ corresponde a los componentes de *V* diferentes de $V_1$, lo que da pie a otros contrafactuales y otras direcciones de descomposición.

**2.2.3. Implementación paramétrica del análisis de descomposición de cambios distributivos.**

El análisis de descomposición anterior puede ser directamente implementado usando representaciones no-paramétricas de las distribuciones involucradas en la descomposición. Sin embargo, es posible obtener una representación paramétrica de las distribuciones condicionales involucradas en los contrafactuales descritos en el apartado anterior. En la práctica esto requiere que la distribución de las variables condicionales (*V*,*W*) sea discreta para así definir grupos de individuos con combinaciones específicas de variables *V* y *W*.

Por razones que tienen que ver con la interpretación de los resultados, en este trabajo se sigue una representación paramétrica de las distribuciones usadas en la definición de contrafactuales en el modelo de descomposición. Analizar cambios en parámetros con significados económicos directos como los retornos a la educación o a las características demográficas de los hogares hacen la discusión de resultados un poco más sencilla que si se consideraran distribuciones condicionales totales.

Una representación paramétrica general de las funciones condicionales $g^{\tau}(y|V,W)$ y $h^{\tau}(V \mid W)$ relaciona *y* con (*V*,*W*) por un lado, y *V* con *W* por el otro. Dichas relaciones pueden expresarse por:

$$y = G\,[V, W, \varepsilon\,;\, \Omega_{\tau}]$$
$$V = H\,[W, \eta;\, \Phi_{\tau}] \qquad (19)$$

Donde $\Omega_{\tau}$ y $\Phi_{\tau}$ son conjuntos de parámetros y $\varepsilon$ y $\eta$ son variables aleatorias. Estas variables aleatorias tienen un papel similar al del término residual en regresiones econométricas, ya que representan la dispersión del ingreso *y* o de las características individuales *V* y *W*. Estas variables aleatorias están distribuidas de acuerdo a las funciones de densidad $\pi^{\tau}(\ )$ y $\mu^{\tau}(\ )$. Las funciones *G* y *H* tienen formas funcionales preestablecidas.

Con esta parametrización, las distribuciones marginales de ingresos en el periodo τ pueden escribirse ahora como:

$$f^{\tau}(y) = \int_{G(V,W,\varepsilon;\Omega_\tau)=y} \pi^{\tau}(\varepsilon) d\varepsilon \left[ \int_{H(W,\eta,\Phi_\tau)=V} \mu^{\tau}(\eta) d\eta \right] \Psi^{\tau}(W) dV dW \qquad (20)$$

Los contrafactuales pueden generarse modificando total o parcialmente el conjunto de parámetros $\Omega_\tau$ y $\Phi_\tau$, las distribuciones $\pi^\tau(\ )$ y $\mu^\tau(\ )$, o la distribución conjunta de características exógenas, $\Psi^\tau(W)$. Estos contrafactuales pueden definirse entonces como:

$$D[\Psi, \pi, \mu; \Omega, \Phi] = \int_{G(V,W,\varepsilon;\Omega)=y} \pi(\varepsilon) d\varepsilon \left[ \int_{H(W,\eta;\Phi)=V} \mu(\eta) d\eta \right] \Psi(W) dV dW \qquad (21)$$

donde cualquiera de las tres distribuciones, $\Psi(\ )$, $\pi(\ )$ y $\mu(\ )$, y los dos conjuntos de parámetros, $\Omega$ y $\Phi$, pueden ser observados en los periodos *t* y *t'*. Por ejemplo, $D[\Psi_t, \pi_t, \mu_t; \Omega_{t'}, \Phi_t]$ podría corresponder a la distribución del ingreso obtenida al aplicar a la población observada en el periodo *t*, los parámetros del modelo de ingreso del periodo *t'*, mientras se mantiene constante la distribución del término residual aleatorio, $\varepsilon$, y todo lo que concierne con las variables *V*. Así, la contribución del cambio en los parámetros de $\Omega_t$ a $\Omega_{t'}$ puede medirse por la diferencia entre $D[\Psi_t, \pi_t, \mu_t; \Omega_{t'}, \Phi_t]$ y $D[\Psi_t, \pi_t, \mu_t; \Omega_t, \Phi_t]$. Desde luego, otras direcciones (paths) de descomposición pueden usarse. Por ejemplo, la comparación podría llevarse a cabo usando la población observada en el periodo *t'* como referencia, en cuyo caso la contribución del cambio en los parámetros Ω podría estar dada por $D[\Psi_{t'}, \pi_{t'}, \mu_{t'}; \Omega_{t'}, \Phi_{t'}] - D[\Psi_{t'}, \pi_{t'}, \mu_{t'}; \Omega_t, \Phi_{t'}]$.

En este esquema paramétrico, el número de direcciones (paths) de descomposición es muy grande. La contribución de cada cambio individual en los parámetros Ω y Φ, en la distribución de los términos residuales $\pi(\ )$ y $\mu(\ )$, y finalmente en la distribución total de las características exógenas $\Psi(\ )$, puede ser evaluada de varias maneras diferentes, dependiendo del valor que se les de a los otros parámetros o a las funciones usadas para las otras distribuciones. En general, es usada una sola dirección de descomposición. Pero es importante comparar los resultados con aquellos obtenidos sobre diferentes direcciones para ver si ellos son muy diferentes, y para comprender las razones de esta diferencia.[12]

[12] Una alternativa podría ser considerar un número grande de direcciones y promediar las contribuciones de cada cambio entre un periodo y otro.