Capítulo 3

Metodología.

3.1. Representación paramétrica de la relación entre el ingreso per cápita de los hogares y las características socio-demográficas de sus miembros.

La metodología utilizada en este trabajo se basa en una representación paramétrica de la forma en que el ingreso per cápita del hogar o los ingresos individuales están relacionados con características socio-demográficas (o ‘dotaciones de activos’) de los individuos y hogares. Este punto de vista es muy similar al método de Oaxaca-Blinder, excepto por dos puntos: a) tiene que ver con la distribución total en lugar de los ingresos medios y b) la representación paramétrica de la generación del ingreso familiar y posiblemente individual es más compleja.

La representación paramétrica utilizada establece que el ingreso de cada hogar es la sumatoria de:

a) los ingresos por salarios de cada miembro,
b) los ingresos por autoempleo y
c) otros ingresos.

La especificación de lo anterior es:

$$Y = \sum_{i=1}^{n} L_i^w w_i + \prod + Y_0 \qquad i = 1,\ldots,n \qquad (22)$$

Donde Y es el ingreso total del hogar, L_i^w son las horas de trabajo en un empleo remunerado salarialmente de cada miembro del hogar *i*, w_i es el salario percibido por cada miembro *i*, $\prod$ es el ingreso en caso de autoempleo, y Y_0 es el ingreso por otras fuentes. En esta especificación, los salarios por hora percibidos están en función de las características individuales de cada miembro (escolaridad y edad) y de la localización geográfica. El ingreso por autoempleo depende de los activos del hogar, de la cantidad de trabajo dedicada a la actividad, de las características de los miembros del hogar dedicados a dicha actividad de autoempleo y de la localización geográfica. La función de oferta laboral está determinada por las características de los miembros del hogar (número de miembros con cierto nivel de escolaridad, edad y género).

Se puede estimar un modelo econométrico estándar con la forma reducida del ingreso del hogar de la ecuación (22):

$$Log\ y_i^\tau = \Omega_\tau X_i^\tau + \varepsilon_i^\tau \qquad (23)$$

donde $Log\ y_i^\tau$ es el logaritmo del ingreso per cápita *y* del hogar *i* en el año τ ajustado para tomar en cuenta cambios en los precios entre regiones, *X* es el vector de características del hogar, Ω es el vector de retornos a tales características y ε es el término de error de la regresión, del cual se asume que $\varepsilon \sim N(0,\sigma^2)$.

Entonces, los contrafactuales *D*( ) definidos anteriormente en la ecuación (21) son fácilmente calculados. Sin la distinción (*V*,*W*), un contrafactual es ahora definido como $D(\chi,\pi;\ \Omega)$, donde χ( ) es la distribución de las características observables *X* y π( ) es la distribución de características inobservables ε. Pasando a una distribución discreta del ingreso $\{y_i\}^\tau = (y_1, y_2, ..., y_{N_\tau})$ en el periodo τ, donde N_τ es el número de observaciones en la muestra disponible en el periodo $\tau = t, t'$, para el año *t* puede establecerse:

$$\{y_i\}^t = D(\chi_t, \pi_t, \Omega_t) \qquad (24)$$

3.2. Descomposición de las contribuciones de los factores microeconómicos en los niveles de desigualdad.

Una vez que se estima el ingreso para cada uno de los años (*t* y *t'*), se calcula la contribución de los cambios en los retornos y en las características observables e inobservables en el cambio total de la distribución del ingreso. Esto se hace simulando la distribución del ingreso que se hubiera observado en *t* si los retornos y las características observables e inobservables hubieran sido los mismos que los obtenidos en *t'* (y viceversa).

Este método de simulación descompone los cambios observados en la distribución del ingreso en "efecto precios", "efecto activos" y "efecto de inobservables". El "efecto precios" captura la contribución que tienen los cambios en los retornos de las características de los hogares, tales como la escolaridad de sus miembros; el "efecto activos" mide la contribución de cambios en el nivel y distribución de tales características de los hogares en su conjunto; por último, el "efecto de inobservables" captura el efecto de cambios en características que no se pueden observar, como las habilidades empresariales de los miembros de cada hogar.

Para medir el "efecto precios" total se cambian simultáneamente todos los coeficientes del vector Ω, con lo que se genera un nuevo vector "simulado" de ingresos per cápita de los hogares. Una vez que se estima la ecuación (23) para *t* y para *t'*, se sustituyen los parámetros obtenidos en la ecuación de *t'* en la ecuación de *t* (y viceversa):

$$Log\ y_i^t = \hat{\Omega}_{t'} X_i^t + \hat{\varepsilon}_i^t \qquad (25)$$

El "efecto precios" total se descompone en los efectos de cambios en retornos específicos, los cuales se calculan reemplazando uno por uno los coeficientes de *t* por aquellos obtenidos en *t'* (y viceversa). Es decir, se repite el procedimiento sustituyendo uno por uno los parámetros de un año en el otro para medir, ceteris paribus, el efecto de cambios en los retornos de cada característica de los hogares (características demográficas, de educación, activos financieros, etc.).

El “efecto de inobservables” se estima modificando la distribución de los residuales, sustituyendo la distribución de los residuales de *t’* con los de *t* y viceversa. Esto se hace escalando los términos de error en un año por la desviación estándar (σ) del otro año[1]:

$$Log\ y_i^t = \hat{\Omega}_t\ X_i^t + \frac{\hat{\sigma}_{t'}}{\hat{\sigma}_t} \hat{\varepsilon}_i^t \qquad (26)$$

Finalmente, el “efecto activos” se calcula con la diferencia entre el cambio total de la distribución del ingreso per cápita de los hogares y los efectos “precios” y “de inobservables” obtenidos con anterioridad.

Tales simulaciones producen un nuevo conjunto de vectores del ingreso per cápita de los hogares. Cada vector simulado de ingreso per cápita representa los ingresos que los hogares pudieron haber recibido bajo diversos supuestos. Con cada vector es posible calcular las medidas de desigualdad y llevar a cabo la descomposición para estimar la contribución de efectos individuales.

El ejercicio de simulación puede resumirse de la siguiente manera: $D(\Omega_t,\ \chi_t,\ \pi_t)$ es la distribución del ingreso de los hogares en el periodo *t*, donde χ es la distribución de las características observables de los hogares, Ω los retornos de tales características y π la distribución de características inobservables. El ejercicio de descomposición consiste en estimar los efectos sobre la distribución del ingreso per cápita de los hogares modificando uno ó más componentes de $D(\cdot)$. El “efecto precios” total se estima cambiando todos los elementos del vector Ω. El efecto individual de retornos específicos se calcula cambiando sólo un elemento del vector Ω. El “efecto de inobservables” se calcula modificando la distribución de los residuales. Por último, el “efecto activos” se calcula como la diferencia entre el cambio total en la distribución del ingreso per cápita de los hogares y los “efectos precios” y “efecto de inobservables” simulados en los pasos anteriores.

[1] Juhn et al. (1993) calcularon residuales simulados basados en el percentil del ingreso actual de un hogar en un año en particular y la distribución de promedio acumulado en el tiempo. Si se asume una distribución normal, entonces esto es equivalente a escalar los términos de error en un año por la desviación estándar del otro año. Este es el procedimiento usado aquí.

El interés radica en explicar el cambio en la distribución del ingreso per cápita de los hogares (ΔD) entre los años t y t':

$$\Delta D = D(\Omega_{t'}, \chi_{t'}, \pi_{t'}) - D(\Omega_t, \chi_t, \pi_t) \qquad (27)$$

Esto puede descomponerse en los tres efectos que se han venido mencionando: “efecto precios”, “efecto de inobservables” (después de haber ajustado los precios), y “efecto activos” (después de haber calculado los efectos “precios” y “de inobservables”). Esto puede establecerse como:

$$\Delta D = [D(\Omega_{t'}, \chi_{t'}, \pi_{t'}) - D(\Omega_t, \chi_{t'}, \pi_{t'})] + [D(\Omega_t, \chi_{t'}, \pi_{t'}) - D(\Omega_t, \chi_{t'}, \pi_t)] + [D(\Omega_t, \chi_{t'}, \pi_t) - D(\Omega_t, \chi_t, \pi_t)] \qquad (28)$$

Simplificando:

$$\Delta D = D_\Omega(\chi_{t'}, \pi_{t'}) + D_\pi(\Omega_t, \chi_{t'}) + D_\chi(\Omega_t, \pi_t) \qquad (29)$$

Donde cada subíndice representa el cambio en la distribución de ingresos como consecuencia de haber cambiado la variable señalada en dichos subíndices. Esta es una descomposición “secuencial” de los efectos precios, inobservables y activos. Sin embargo, esta descomposición no mantiene las condiciones finales (correspondientes a t') constantes en cada paso de la simulación. La descomposición “secuencial” mide:

a) el “efecto precios” usando las características observables e inobservables de los hogares en el periodo t',
b) el “efecto de inobservables” usando los retornos en t y las características observables de los hogares en t', y
c) el “efecto activos” usando los retornos y las características inobservables en t.

Para mantener las condiciones finales constantes y medir todos los efectos usando variables del periodo *t'*, se aplica una transformación reacomodando los términos, obteniendo:

$$\Delta D = D_{\Omega}(\chi_{t'},\pi_{t'}) + D_{\pi}(\Omega_{t'}, \chi_{t'}) + D_{\chi}(\Omega_{t},\pi_{t}) + [D_{\pi}(\Omega_{t}, \chi_{t'}) - D_{\pi}(\Omega_{t'}, \chi_{t'})] \quad (30)$$

"ef. precios" + "ef. de inobservables" + "ef. activos" + remanente

O alternativamente, para mantener las condiciones iniciales (en *t*) constantes:

$$-\Delta D = D_{\Omega}(\chi_{t},\pi_{t}) + D_{\pi}(\Omega_{t}, \chi_{t}) + D_{\chi}(\Omega_{t'},\pi_{t'}) + [D_{\pi}(\Omega_{t'}, \chi_{t}) - D_{\pi}(\Omega_{t}, \chi_{t})] \quad (31)$$

Las ecuaciones (30) y (31) indican que el cambio total en la distribución del ingreso per cápita puede expresarse como la suma de los efectos "precios" y "de inobservables" – dadas las condiciones finales (iniciales) – más el "efecto activos" – dadas las condiciones iniciales (finales) – más un remanente. Los términos remanentes son la interacción entre diferentes factores siendo simulados, es decir, muestran que el efecto combinado de modificar dos ó más factores al mismo tiempo – dígase retornos y características de los hogares – no es igual a la suma de los componentes por separado, que en este caso son el efecto de cambiar retornos manteniendo constantes las características de los hogares y el efecto de cambiar tales características manteniendo los retornos constantes.[2]

Si se establece el supuesto razonable de monotonicidad de la descomposición en cambios en retornos, características de los hogares y factores inobservables, los resultados de las descomposiciones (30) y (31) representan los límites superior e inferior de las estimaciones. Tomando este supuesto, el análisis de los resultados se basa en el promedio de los límites superior e inferior.[3]

[2] El mismo procedimiento se aplica para descomponer el "efecto precios" total en efectos precios específicos. El efecto combinado de modificar dos o más retornos al mismo tiempo – por ejemplo los retornos de características de los hogares y retornos regionales – no será igual a la suma de modificar cada retorno manteniendo los demás constantes. Por lo tanto, la descomposición del "efecto precios" total en sus componentes incluye un término remanente que refleja el hecho de que esa descomposición no fue hecha de manera secuencial para mantener las condiciones iniciales (finales) constantes.

[3] Esto se refiere a que si, por ejemplo, tomamos el periodo 1984-2002, se calculan primero las contribuciones tomando como año base 1984 (límite inferior), y posteriormente se obtienen las contribuciones tomando como año base 2002 (límite superior), para así poder calcular las contribuciones promedio entre 1984 y 2002.