

# *Las características educativas de los emigrantes mexicanos a Estados Unidos*

ALFREDO CUECUECHA MENDOZA<sup>1</sup>

- **Resumen:** Este artículo estudia la auto-selección en características educativas de mexicanos que emigran a Estados Unidos. Esto se hace para determinar la auto-selección educativa, los cambios en la auto-selección en la década de los noventa, así como el cambio en la auto-selección en términos de la distribución de salarios. El artículo aplica técnicas econométricas que controlan por el sesgo que genera el uso de muestras endógenas. Los cuatro principales resultados del artículo son: primero, la autoselección en los noventa era intermedia; segundo, la autoselección en el 2000 pasó a los extremos de cero años y 17 años o más; tercero, dichos grupos son los que experimentaron el crecimiento más importante durante los noventa y, cuarto, es que a pesar de los cambios en la distribución de educación de mexicanos en Estados Unidos, la auto-selección en términos de la distribución salarial cambió en los noventa fundamentalmente hacia una selección intermedia.
- **Abstract:** This paper studies the self selection in educational characteristics of Mexicans migrating to the US. This is done to determine educational self-selection, changes in self-selection in the nineties, as well as the change in the self-selection in terms of the wage distribution. The paper applies econometric techniques that control for bias by the use of endogenous samples. There are four key findings: first, self selection in the nineties was intermediate; second, self selection in 2000 was in the extremes of the education distribution: at zero years and at more than 16 years of education; third, the largest changes in migration probability were experienced by those two groups; and fourth, despite the changes in migration probabilities, the change in self selection in terms of the wage distribution is a change towards intermediate selection.

---

<sup>1</sup> Universidad Iberoamericana Puebla. alfredo.cuecuecha@iberopuebla.edu.mx. Agradezco los comentarios de los dictaminadores anónimos que ayudaron a mejorar significativamente el documento. Cualquier error que persista es de mi responsabilidad.

- **Palabras clave:** Migración, muestreo basado en decisiones de personas, educación.
- **Clasificación JEL:** F22, C25, C81, I29
- Fecha de recepción: 04/03/2010                      Aceptación: 25/08/2010

### ■ *Introducción*

El objetivo de este artículo es comparar las características educativas de los emigrantes mexicanos a Estados Unidos con las características educativas de los mexicanos que permanecen en México. Este objetivo es importante debido a que permite obtener información sobre el impacto de la emigración para grupos educativos específicos.

El impacto específico en grupos educativos es importante para México pues, por ejemplo, permite determinar si la emigración genera fuga de cerebros en el país, lo cual podría determinarse si existe una relación positiva entre la emigración y la educación (Lowell y Pederzini, 2010).

Un segundo objetivo es determinar el impacto agregado de la emigración. Este objetivo es importante de obtener por diferentes razones. Para los países receptores, les permite establecer el posible impacto fiscal de la emigración. Para los países expulsores, les permite establecer el posible impacto de la emigración sobre la desigualdad salarial.

Bajo el supuesto de que los individuos con altas habilidades educativas también tendrán altas habilidades no observadas (Chiswick, 1986) es posible determinar el impacto agregado de la emigración conociendo las características educativas de los emigrantes mexicanos. Por ejemplo, si los individuos son seleccionados negativamente en características no observadas, dichos individuos representarán una carga fiscal para el país receptor, pues utilizarán con mayor frecuencia los servicios de beneficencia pública (Borjas, 1987).

En caso de que existan mercados incompletos de créditos, la validez de este argumento no es necesariamente cierta, pues el nivel de educación ya no está necesariamente correlacionado con las habilidades no observadas (Mackenzie y Rapoport, 2007). Así mismo, la existencia de redes sociales migratorias puede alterar estos patrones (Massey y Espinoza, 1997; Mackenzie y Rapoport, 2007a). Por ello, además de determinar la selección en educación, este artículo muestra la selección en características observables en términos de las distribu-

ciones salariales, usando un método propuesto por Chiquiar y Hanson (2005).<sup>2</sup>

Los datos ideales para realizar este estudio deberían ser una muestra aleatoria representativa de los mexicanos que emigran a Estados Unidos y una muestra aleatoria representativa de los mexicanos que permanecen en México. Este artículo utiliza las muestras públicas de los censos de México y Estados Unidos para los años 1990 y 2000. Ambas representan bases de datos estratificadas y su combinación representa un ejemplo de un muestreo basado en las decisiones de los individuos estudiados. A este tipo de muestras también se les conoce como muestras endógenas (Amemiya, 1985) y este término se utilizará a partir de ahora.

La muestra estratificada que se utiliza en este artículo se dice que es una muestra endógena debido a que los individuos han decidido su lugar de residencia y debido a que los individuos se han muestreado con diferentes probabilidades, dependiendo del lugar de residencia y del año del censo al cual correspondan. La importancia de este hecho se puede mostrar mediante un ejemplo sencillo. Suponga que tenemos una muestra de 10 mexicanos en Estados Unidos y de 20 mexicanos en México. Si no tomamos en consideración la tasa a la cual los hemos muestreado, podemos concluir erróneamente que 30% de la población mexicana se encuentra en Estados Unidos.

La literatura ha propuesto ya soluciones para el problema de las muestras endógenas. Manski y Lerman (1977) proponen el uso de estimadores ponderados que permiten resolver el problema de las muestras endógenas. Estos métodos se conocen como los *Weighted Exogenous Sampling Estimators (WESML)*. Imbens (1992) propone un método que utiliza el método generalizado de momentos (*GMM*) que genera estimadores consistentes y eficientes.<sup>3</sup>

Debido a que la determinación de los impactos sobre grupos educativos específicos y a que la determinación del impacto agregado de la emigración dependen crucialmente del uso de técnicas econométricas que corrijan el sesgo generado por usar muestras endógenas, un tercer objetivo de este artículo es determinar la técnica econométrica más apropiada para los datos mexicanos. Este artículo es el primero en aplicar tanto las técni-

<sup>2</sup> Es importante mencionar que existen otros autores para quienes los verdaderos efectos de la migración sobre el mercado receptor deben medirse directamente en los impactos que la emigración tiene sobre los mercados locales de trabajo en los Estados Unidos. La explicación es que los métodos contrafactuales no pueden controlar por posibles efectos de equilibrio general que la emigración genere. Véase Card y Lewis (2007) para un ejemplo de esta literatura.

<sup>3</sup> Otros artículos han aplicado su propia versión de las técnicas WESML, por ejemplo Bover y Arellano (2002), para estudiar la migración interna en España.

cas de *WESML* como de *GMM* para estudiar las características educativas de los emigrantes mexicanos a Estados Unidos, usando datos de los censos de México y Estados Unidos.<sup>4</sup> El artículo muestra que la aplicación de la técnica *GMM* supera a la técnica *WESML* debido a que la primera puede predecir mejor la tasa de emigración agregada. Así mismo, el artículo muestra la importancia del uso de estas técnicas para la estimación de densidades salariales no paramétricas, pues se muestra que el no utilizarlas genera una subestimación del efecto agregado de la emigración.

Existe una literatura creciente sobre el tema de la selección de emigrantes mexicanos. Algunos han encontrado selección negativa (Taylor, 1986; Borjas, 1987; Orrenious, 1999; Ibararán y Lubotsky, 2007; Fernández-Huertas, 2008; Kaestner y Malamud, 2010). Otros autores han encontrado selección intermedia o positiva (Chiquiar y Hanson, 2005; Orrenious y Zavodny, 2005; Mishra, 2007; Cuecuecha, 2008; Lowell, Passel y Pederzini, 2008). Finalmente, existe un artículo que ha encontrado que la mayor emigración ocurre en los extremos de la distribución de educación (Cuecuecha, González y Pederzini, 2010).

Este artículo también toma en consideración la existencia de subconteo de mexicanos en el censo de Estados Unidos. De acuerdo a Passel (2004, 2005a, 2005b); Passel, Capps y Fix (2004); Passel y Fix (2001); y Passel, Van Hook y Bean (2004), la tasa de subconteo de mexicanos en Estados Unidos alcanza un 10% en el censo del año 2000, mientras que la tasa de subconteo en el censo de 1990 se sitúa entre el 15% y el 20%. En este estudio se asume que el subconteo en el censo mexicano es de cero.<sup>5</sup> El subconteo genera estimaciones sesgadas en el caso de los estimadores *WESML* y potencialmente en los estimadores *GMM*.

El artículo presenta un método novedoso para resolver el problema. Este método consiste en aumentar el número de observaciones haciendo supuestos acerca de la distribución de educación de los subcontados.<sup>6</sup>

---

4 Cuecuecha (2008) usa la metodología *WESML* y *GMM* pero usa los datos de la ENIGH 1994 y CPS 1994. La ENIGH 1994 no cuenta con suficientes observaciones en las colas de la distribución de educación.

5 Ibararán y Lubotsky (2007) argumentan que el subconteo puede ocurrir en el censo mexicano si los individuos no responden a las preguntas de edad, educación, ingreso o si no contestan el censo. Para dar una idea de la magnitud del problema, la muestra pública del censo del año 2000 contabiliza .35% de individuos que no tienen la información sobre la edad y 3.17% de observaciones para hombres entre 25 y 28 años de edad que no tienen la información sobre educación. En la muestra pública del censo del año 1990, todas las observaciones tienen la información de edad y de educación. Dado que la solución de estos problemas requiere de técnicas econométricas diferentes a las mencionadas en este artículo y dado que la prevalencia del problema es mucho menor a la importancia del subconteo en los censos de Estados Unidos, decidimos ignorar el problema. Una estrategia similar ha sido seguida por Chiquiar y Hanson (2005).

6 Estos supuestos son: a) asumir que los subcontados se auto-seleccionan en la misma forma

Se muestra que, en términos generales, incluir a los subcontados ayuda a determinar los niveles de migración para 1990 y 2000, pero que los cambios a través del tiempo y la estimación de las densidades salariales no se ven afectadas en gran medida por el subconteo.

El resto del artículo está organizado como sigue: la segunda parte presenta la metodología empleada en el estudio. La tercera sección presenta los datos empleados en el artículo y explica a detalle la metodología para tratar el problema de los subcontados. La cuarta sección presenta los resultados de la estimación y análisis de robustez de los resultados. La quinta sección muestra las implicaciones de nuestros resultados para la elaboración de distribuciones no paramétricas de salarios. La sexta sección concluye el documento.

### ■ *Metodología econométrica*

#### *Técnicas para Muestras Endógenas*

Sea la probabilidad de elegir Estados Unidos, condicionada en el vector de características  $X$ ,  $P(US=1|X)$ , definida como:

$$(1) \quad P(US=1|X) = P(\delta'X)$$

Donde  $\delta$  mide la contribución de la característica  $X_k$  a la función lineal  $\delta'X$ ,  $h$  se define como la frecuencia con la cual los individuos en Estados Unidos están siendo muestreados y  $q$  como la verdadera tasa de emigración. Si  $R(X)$  es la función de distribución acumulada de las características de los individuos y  $r(x)$  es la función de densidad de las características, entonces  $q = \int P(\delta'X) dR(X)$ . Notar que si la muestra fuera aleatoria, esta sería toda la información que se necesita para construir la función de verosimilitud. Para tomar en consideración la tasa de muestreo, la probabilidad de encontrar un mexicano en Estados Unidos, dadas sus características  $X$ , está dada por:

$$(2) \quad g(US, X) = \frac{h}{q} P(\delta'X)r(X)$$

donde  $\frac{P(\delta'X)}{q}$  representa la probabilidad condicional de encontrar a un

---

que el resto de la población de mexicanos en Estados Unidos y b) que son autoseleccionados de una forma diferente al del resto de la población. Se asume para el segundo supuesto que todos los subcontados son ilegales. La distribución de educación de los emigrantes ilegales se obtiene de la base de datos MMP 93. Para mayor información sobre el MMP 93 consulte <http://mmp.opr.princeton.edu/webnews/news-en.aspx>.

mexicano en Estados Unidos, condicionada en estar en la muestra de mexicanos en Estados Unidos. La ecuación 2 sintetiza la razón por la cual una regresión que no toma en consideración la tasa de muestreo es sesgada e inconsistente: dado que la frecuencia de muestreo puede ser diferente de la propensión con la cual los individuos deciden migrar, una regresión de tipo probit o logit que no considere el muestreo endógeno especificará incorrectamente la función de verosimilitud.

### Estimación WESML

Manski y Lerman (1977) argumentan que si los parámetros  $(h, q)$  son conocidos, es posible multiplicar cada observación en la muestra de Estados Unidos por la razón  $w(US) = q/h$ , y a todas las observaciones en México por  $w(MX) = (1-q)/(1-h)$ .<sup>7</sup> Su estimador se llama el *Weighted Exogenous Sampling Maximum Likelihood (WESML)*. La función de verosimilitud está dada por (Amemiya, 1985: 322):

$$(3) \quad L_{WESML}(\delta) = \sum_i I(US)_i w(US) \ln P(\delta' X_i) + \sum_i I(MX)_i w(MX) \ln(1 - P(\delta' X_i))$$

donde  $I(j)_i$  es una variable indicador de país, donde el país  $j=US, MX$ , para el individuo  $i$ . Es importante mencionar que la consistencia de los estimadores *WESML* depende fuertemente del conocimiento que se tenga de los parámetros  $(h, q)$ . Si estos parámetros se conocen con error, el estimador *WESML* es inconsistente.

Siguiendo a Amemiya (1985), el estimador *WESML* es asintóticamente normal, consistente y con varianza dada por:

$$(3a) \quad \Gamma^{WESML} = A^{-1} \Lambda A^{-1}$$

donde:

$$(3b) \quad A = -E w(j) S(\delta) S(\delta)' \quad S(\delta) = \frac{\partial L(\delta)}{\partial \delta}$$

$$(3c) \quad \Lambda = E w(j)^2 S(\delta) S(\delta)'$$

<sup>7</sup> Es importante mencionar que, en la estimación, los pesos para cada observación también son multiplicados por los factores de expansión proveídos por los censos mexicanos y de EU. Esto es sugerido por DiNardo, Fortin y Lemieux (1996), quienes crearon el método que Chiquiar y Hanson (2005) adaptaron al caso de la emigración de mexicanos.

*Estimación GMM*

El estimador *WESML* es consistente, pero ineficiente, por lo cual otros estimadores han sido propuestos. Este problema se suma al hecho de que, en la práctica, un conocimiento perfecto de los parámetros ( $h, q$ ) puede no estar disponible. Imbens (1992) propone un estimador en el cual el conocimiento perfecto de los parámetros ( $h, q$ ) no es necesario para desarrollar el método, pues forman parte del conjunto de parámetros a estimar. Además, el estimador se construye buscando minimizar la varianza para generar un estimador consistente y eficiente. Imbens (1992) propone el uso de tres momentos generados a partir de tres condiciones que están dadas por las siguientes ecuaciones:<sup>8</sup>

$$(4) \quad \begin{aligned} \Psi_1(j) &= h(j) - I(j) \\ \Psi_2(j) &= q(j) - \frac{P(j|X, \delta)}{\Delta} \end{aligned}$$

$$\Psi_3(j) = P_\delta(j|X, \delta) \frac{1}{P(j|X, \delta)} - P_\delta(\delta'X) \frac{\left[ \frac{h}{q} - \frac{1-h}{1-q} \right]}{\Delta}$$

donde:

$$(5) \quad \Delta = \frac{h}{q} P(\delta'X) + \frac{1-h}{1-q} [1 - P(\delta'X)]$$

para individuos ubicados en el país  $j=MX, US$ , y  $h(MX)=1-h$ , mientras que  $q(MX)=1-q$ . Del mismo modo,  $P(US|X, \delta)$  representa la probabilidad condicional de migrar para individuos en la muestra de mexicanos en Estados Unidos, mientras que  $P(MX|X, \delta)$  representa el complemento de la probabilidad de migrar en la muestra de mexicanos en México. Finalmente,  $P_\delta(j|X, \delta)$  representa el vector gradiente de la probabilidad de ser observado en el país  $j$ . Notar que el segundo término  $\Psi_3$  es similar para todas las observaciones de la muestra.<sup>9</sup>

El vector de parámetros a estimar en el modelo es  $\delta^* = [h, q, \delta']$  y está definido como sigue<sup>10</sup>:

<sup>8</sup> Vea Imbens (1992) para una mayor explicación sobre la derivación de estos tres momentos.

<sup>9</sup> Para mayor detalle en la interpretación de los momentos consulte Cuecuecha (2008).

<sup>10</sup> Notar que en la estimación de la ecuación 6 cada vector que corresponde a un individuo dado es multiplicado por los factores de expansión proporcionados por los censos de México y Estados Unidos.

$$\begin{aligned} \delta^* &= \arg \min R_n(\delta^*) \\ (6) \quad R_n(\delta^*) &= \frac{1}{n} \sum_i \Psi(\delta^*, X_i, j)' C_n \frac{1}{n} \sum_i \Psi(\delta^*, X_i, j) \end{aligned}$$

Finalmente, la matriz óptima de ponderación  $C_n$  está dada por:

$$(7) \quad C_n = \left[ \frac{1}{n} \sum_i \Psi(\delta^*, X_i, j) \Psi(\delta^*, X_i, j)' \right]^{-1}$$

Defina las matrices  $\Delta_0$  y  $\Gamma_0$  como sigue:

$$\begin{aligned} (8) \quad \Delta_0 &= E \Psi(\delta^*, X_i, j) \Psi(\delta^*, X_i, j)' \\ \Gamma_0 &= E \frac{\partial \Psi(\delta^*, X_i, j)}{\partial \delta^*} \end{aligned}$$

Dado el supuesto de que el modelo a estimar es un probit y otras condiciones de regularidad,<sup>11</sup> Imbens (1992) demuestra que  $\delta^*$  converge casi seguramente a un vector con distribución Normal:

$$(9) \quad n^{1/2}(\hat{\delta}^* - \delta^*) = N(0, \Gamma_0^{-1} \Delta_0 \Gamma_0^{-1})$$

## ■ Datos

### *Datos agregados*

El uso de las técnicas *WESML* requiere una estimación de la tasa de emigración agregada (Imbens y Lancaster, 1994). El cuadro 1 nos muestra las estimaciones que se utilizan para la tasa de migración agregada, considerando el subconteo y sin considerarlo. Los datos sobre la población mexicana en México en 1990 y 2000 se obtienen de INEGI (2010). Los datos para la población mexicana en Estados Unidos se obtienen de Chiquiar y Hanson (2005). Finalmente, se asume una tasa de subconteo del 15% para 1990 y del 10% para el año 2000 (Passel, 2004; 2005a; 2005b; Passel, Capps y Fix, 2004; Passel y Fix, 2001; Passel, Van Hook y Bean, 2004). Este cuadro nos muestra que la tasa de migración agregada en 1990 oscila entre el 5.02% y el 5.74% de la población de mexi-

<sup>11</sup> Para mayor detalle en las condiciones de regularidad consulte Imbens (1992).



**Cuadro 1**  
**Tasas de emigración, 1990-2000**

	Mexicanos en México (millones)	Mexicanos en EU sin subcontados (millones)	Mexicanos en EU con los subcontados (millones)	Tasa de Migración sin subconteo (%)	Tasa de Migración con subconteo (%)
1990	81.25*	4.30**	4.95**	5.02	5.74
2000	97.48*	9.18**	10.10**	8.60	9.38

Fuentes: \* INEGI (2010) e IPUMS (2010) \*\* Chiquiar y Hanson (2005).

Nota: Se utiliza una tasa de subconteo del 15% para 1990 y de 10% para el año 2000. La fuente para estas tasas de subconteo son Passel (2004, 2005a, 2005b); Passel, Capps y Fix (2004); Passel y Fix (2001); Passel, Van Hook y Bean (2004).

canos, mientras que para el año 2000, la tasa de emigración oscila entre el 8.60% y el 9.38% de la población de mexicanos.<sup>12</sup>

### *Datos individuales*

Los datos para mexicanos en México en el año de 1990 provienen de la muestra pública que representa una muestra aleatoria del 1% del total de los archivos individuales del censo de 1990. Se escogen únicamente mexicanos entre las edades de 24 a 58 años en 1990, lo cual genera un total de 127,932 observaciones (INEGI, 2010). Los datos para los mexicanos en Estados Unidos provienen de la muestra del 5% del censo de Estados Unidos del año 1990 (IPUMS, 2010). La muestra sólo incluye individuos que entraron a Estados Unidos cuando tenían 17 años o más para reducir el número de individuos que obtuvieron su educación en Estados Unidos.<sup>13</sup> Al igual que para la muestra de mexicanos en México sólo se consideran individuos que tenían entre 24 y 58 años en 1990, lo cual genera un total de 50,153 observaciones.

Los datos para el año 2000 se obtienen de las muestras públicas de los censos de México y Estados Unidos. En el caso de México, ésta es una muestra del 9.1% del total de los registros individuales del censo del año 2000. En el caso de Estados Unidos, ésta es una muestra del 5% de los registros individuales del censo del año 2000. El total de observacio-

<sup>12</sup> Estas tasas se utilizarán en la implementación de las metodologías WESML y GMM. De hecho, servirán como base para decidir sobre cuál de los dos métodos se adapta mejor a los datos mexicanos.

<sup>13</sup> El apéndice explica cómo se determinó el año de entrada a Estados Unidos. De acuerdo a Lowell, Passel y Pederzini (2008) hay evidencia de que los mexicanos que entran antes de los 17 años tienen mayores posibilidades de estudiar y asimilarse a Estados Unidos.

nes para México es de 1,831,978 individuos y para Estados Unidos es de 344,518 individuos.

El cuadro 2 presenta las características de la muestra analizada en este artículo. El cuadro muestra que la educación promedio de los mexicanos en Estados Unidos es mayor a la educación promedio de los mexicanos en México y que los mexicanos en Estados Unidos también son más jóvenes. Si solamente se analizan migrantes que entraron a Estados Unidos durante los cinco años previos a cada censo, estos resultados son más fuertes<sup>14</sup>.

**Cuadro 2**  
Características de la Muestra, hombres mexicanos de 24 a 58 años. <sup>a</sup>

	1990			2000		
	Mexicanos en México <sup>b</sup>	Estados Unidos <sup>c</sup> Todos	Emigrantes Recientes <sup>f</sup>	México <sup>d</sup>	Estados Unidos <sup>e</sup> Todos	Emigrantes Recientes <sup>f</sup>
Educación	6.61 (5.08)	7.11 (4.72)	7.76 (4.95)	7.66 (4.86)	8.16 (4.54)	8.61 (4.46)
Edad	37.78 (9.3)	37.52 (8.85)	32.82 (7.86)	37.93 (9.12)	37.79 (8.80)	33.10 (7.72)
Observaciones	127,932	50,146	11,959	1,831,978	344,518	82,134

Fuentes: \* INEGI (2010) e IPUMS (2010).

Notas:

<sup>a</sup>Todos los individuos en la muestra de mexicanos en Estados Unidos emigraron a dicho país después de los 17 años.

<sup>b</sup>Muestra del 1% del censo de México 1990.

<sup>c</sup>Muestra del 5% del censo de Estados Unidos 1990.

<sup>d</sup>Muestra del 9.1% del censo México 2000.

<sup>e</sup>Muestra del 5% del censo 2000 de Estados Unidos.

<sup>f</sup>Sólo individuos que entraron cinco años antes del censo en cuestión.

En este artículo, una pieza importante de información es la tasa de muestreo de mexicanos en Estados Unidos. En una muestra aleatoria, este número sería simplemente el número de observaciones que se tienen en la muestra de mexicanos en Estados Unidos dividido entre el total de observaciones en la muestra. Dado que en todos los censos se cuenta con los factores de inflación, se utilizan dichos factores para sa-

<sup>14</sup> En el apéndice se explica cómo se miden los años de educación en los diferentes conjuntos de datos analizados y la distribución de educación para mexicanos en México y la distribución de educación para mexicanos en Estados Unidos para los años analizados.

ber el número de individuos que cada observación representa y de ese modo saber la verdadera tasa de muestreo. Defina a  $TI$  como la suma de los factores de inflación tanto en México como en Estados Unidos. Esta suma representa el total de observaciones con las cuales se cuenta. La tasa de muestreo de mexicanos en Estados Unidos es simplemente la suma de todos los factores de inflación de dichas observaciones dividida entre  $TI$ . La tasa de muestreo para mexicanos en México es el complemento de esta tasa de muestreo. La tasa de muestreo de mexicanos en Estados Unidos es de .3301 para 1990 y de .5222 para el 2000.<sup>15</sup>

### *Los subcontados*

Passel (2004, 2005a, 2005b); Passel, Capps y Fix (2004); Passel y Fix (2001); y Passel, Van Hook y Bean (2004) han estimado la tasa de subconteo en alrededor de 15% para el censo de 1990 y de 10% para el año 2000; este subconteo se debe, entre otras causas, a que hay un número importante de migrantes ilegales que no son contados. El problema es especialmente severo para el estimador *WESML* debido a que los pesos usados en el estimador representan erróneamente a las observaciones. En principio, el estimador *GMM* puede resolver el problema porque los ponderadores son un parámetro más a estimar. Sin embargo, si los subcontados son autoseleccionados, el estimador *GMM* también será sesgado.

El enfoque que se toma en este artículo para resolver el problema es aumentar la muestra de mexicanos en Estados Unidos en proporción a la tasa de subconteo estimada por Passel (2004, 2005a, 2005b); Passel, Capps y Fix (2004); Passel y Fix (2001); y Passel, Van Hook y Bean (2004). Esto implica que para 1990 se aumente el número de observaciones en 15%, mientras que para el año 2000 esto implica un incremento del 10% en el número de observaciones. Las tasas de muestreo y de migración se ajustan a los nuevos datos.<sup>16</sup>

<sup>15</sup> Los factores de inflación que se utilizan en la estimación se ajustan usando un método similar. Se suman los factores de inflación por país y por año y después cada factor de inflación se divide entre el total de la suma de factores de inflación. Este número se invierte y se obtienen los nuevos factores de expansión. De esta forma se garantiza que la suma de los inversos de factores de inflación sumen a uno. Es recomendable hacer este ajuste puesto que los factores de inflación originales fueron calculados sobre la muestra total que se tomó del Censo correspondiente y este artículo sólo utiliza sub-muestras de la muestra original.

<sup>16</sup> Las tasas de muestreo y de migración se ajustan a la inclusión de los sub-contados para garantizar que los factores de expansión representen a los subcontados correctamente. El ajuste se hace siguiendo el procedimiento descrito en el pie de página 15. Es una corrección parcial para tomar en cuenta a los subcontados, pues se necesitaría tener una encuesta representativa de los subcontados para calcular su número y sus características.

Un problema que tiene este enfoque para el estudio de las características educativas es que no sabemos cuál es la distribución de educación entre los subcontados. Se hace el supuesto de que todos los subcontados son emigrantes ilegales, para poder usar las encuestas existentes sobre los ilegales en Estados Unidos provenientes de la encuesta MMP93 y tener una distribución de educación para los subcontados. También se hace este supuesto con la intención de generar la cota máxima de selección negativa que la inclusión de los subcontados puede ocasionar en nuestros resultados, pues los emigrantes ilegales tienen en promedio menor educación que los mexicanos que se encuentran en cualquiera de los censos de Estados Unidos para 1990 y 2000. Para aproximar la distribución de educación de ilegales en 1990, se toma una muestra de los años 1982 a 1992 para representar la distribución para 1990; mientras que se toma una muestra de los años 1992 a 2002 para representar la distribución para 2000.<sup>17</sup> El apéndice da mayores detalles sobre la distribución de educación en el MMP 93.

Con base en esta distribución de educación se obtiene una muestra estratificada de observaciones de los censos Estados Unidos 1990 y Estados Unidos 2000. Estas observaciones se añaden a la muestra original. A la muestra combinada de datos originales y datos estratificados se le llamará a partir de ahora la “Muestra Aumentada Estratificada” (MAE).

Además, para medir la robustez del método se aumentan los datos haciendo el supuesto de que el censo de Estados Unidos representa correctamente a los subcontados. Es decir, esto asume que los subcontados se auto-seleccionan en la misma manera que el resto de los emigrantes mexicanos a Estados Unidos.<sup>18</sup> En este caso se aumenta la muestra original tomando una muestra aleatoria de los censos Estados Unidos 1990 y 2000. A esta muestra combinada de datos originales y datos aleatorios se le llamará a partir de ahora la “Muestra Aumentada Aleatoria” (MAA).

La columna 1 en el cuadro 3 muestra la distribución de educación de la muestra aleatoria tomada del censo de 1990 de Estados Unidos. La columna 3 en el cuadro 3 muestra la distribución de educación para la muestra aleatoria tomada del censo del 2000 de Estados Unidos. El total de observaciones añadido en cada año es diferente, pues las tasas de subconteo son diferentes para cada año, como se explicó anteriormente, y debido a que los tamaños de muestra de cada año son también diferentes.

<sup>17</sup> Orrenius (1999), Orrenious y Zavodny (2005) y Chiquiar y Hanson (2005) también utilizan la encuesta MMP 93 para generar la distribución de educación de los emigrantes ilegales mexicanos.

<sup>18</sup> Passel (2004) también hace este supuesto para corregir sus ponderadores por subconteo.

El cuadro 3 muestra el número de observaciones que se aumentó a los datos originales usando la muestra estratificada. La columna 2 del cuadro 3 muestra la distribución de educación para los subcontados en 1990 y la columna 4 del cuadro 3 muestra la distribución de educación para los subcontados en el año 2000.

Los resultados de la utilización de las muestras MAE y MAA se muestran en la sección de resultados a continuación.

**Cuadro 3**  
Número de observaciones adicionadas a la muestra  
y su distribución de educación por año de censo y Método  
de selección de Muestra Aumentada

Categoría educativa	1990. Incremento del 15% en el número de observaciones		2000. Incremento del 10% en el número de observaciones	
	Muestra aleatoria	Muestra estratificada	Muestra aleatoria	Muestra estratificada
	(%)	(%)	(%)	(%)
Sin educación	13.56	18.79	9.91	11.39
1 a 4 años de educación	16.83	38.36	8.26	32.52
5 a 8 años	32.28	26.37	31.72	33.62
9 a 12 años	27.28	12.56	36.93	17.03
12 a 16 años	8.59	6.29	11.35	3.19
17+	1.45	.62	1.83	2.25
Observaciones	7521	7521	34451	34451

Fuentes: INEGI (2010), IPUMS (2010) y MMP 93.

Notas: Los datos para mexicanos en Estados Unidos en 1990 se aumentan en 15% para tratar de representar a los individuos subcontados en dicho año. Para el año 2000 se aumentan en 10% para representar a los individuos subcontados en dicho año. Estas tasas de subconteo se obtienen de Passel (2004, 2005a, 2005b); Passel, Capps y Fix (2004); Passel y Fix (2001); Passel, Van Hook y Bean (2004). La columna de muestra aleatoria representa el ejercicio de aumentar la muestra bajo el supuesto de que los subcontados no se autoseleccionan y, por lo tanto, una muestra aleatoria de los mexicanos en los censos en EU representa correctamente sus características educativas. Las columnas de muestra estratificada representan el ejercicio de aumentar la muestra bajo el supuesto de que todos los subcontados son ilegales y, por lo tanto, se requiere una muestra estratificada que incluya observaciones de acuerdo a información adicional sobre la distribución de educación de los ilegales. La estratificación de 1990 se hace usando la distribución de educación para emigrantes ilegales para el periodo 1982-1992 que se obtiene de la MMP 93. La estratificación de 2000 se hace usando la distribución de educación para emigrantes mexicanos ilegales para el periodo 1992-2002 que se obtiene de la MMP 93. Mayor información sobre estas distribuciones se encuentra en el cuadro A.2.

## ■ *Estimación*

### *Resultados para 1990*

El cuadro 4 muestra los resultados para las estimaciones hechas para el año 1990. Todos los coeficientes son significativos al 1%. Los coeficientes de los estimadores *WESML* y *GMM* son estadísticamente diferentes de acuerdo a la prueba de Hausman (1978) que también se muestra en el cuadro 4. Es evidente que la diferencia básica se da en el intercepto. Esta diferencia genera que los estimadores *GMM* puedan generar una tasa de emigración muy cercana a la observada, mientras que los estimadores *WESML* generan una tasa de emigración muy lejana a la observada. Éste es un problema ya reconocido en la literatura para los estimadores *WESML* (Manski y Lerman, 1977). Esto nos lleva a preferir las estimaciones *GMM*.

Para interpretar mejor los resultados se presenta la gráfica 1, en la cual se muestra que el polinomio de edad estimado implica que la probabilidad de emigrar encuentra su máximo a los 42 años. Por su parte, el polinomio de cuarto grado en educación muestra una gráfica no monótona que comienza con un pico en los cero años de educación y presenta un valle en los 2.5 años de educación, para continuar con una etapa que crece hasta los 11 años de educación para posteriormente volver a reducirse.<sup>19</sup> La gráfica 2 nos muestra que la inclusión de los subcontados no altera cualitativamente los resultados obtenidos para 1990.

La gráfica 3 compara las tasas de emigración promedio por grupos de educación. En ese sentido sintetiza la información de las gráficas 1 y 2. La gráfica muestra que todos los niveles de educación por encima de los 5 años de educación muestran una tasa de emigración mayor a la de los grupos con 4 años de educación o menos cuando no utilizamos la muestra MAE y, por lo tanto, implicarían una selección en características educativas, tanto intermedia como positiva. Cuando incluimos la muestra MAE, el grupo de educación con cero años de educación aparece con un nivel de emigración tan alto como el grupo de educación con 17 o más años de educación, lo cual nos mostraría que en el agregado y al incluir a los subcontados, la emigración tiende más a mostrar una selección intermedia en características educativas.

### *Resultados para el 2000*

El cuadro 5 muestra los resultados para el año 2000. El cuadro nos muestra que todos los coeficientes son estadísticamente significativos y

<sup>19</sup> Se realizó una prueba F para determinar si era necesario incluir un quinto grado en el polinomio. Sin embargo, no se encontró evidencia estadística de que el quinto grado fuera necesario.

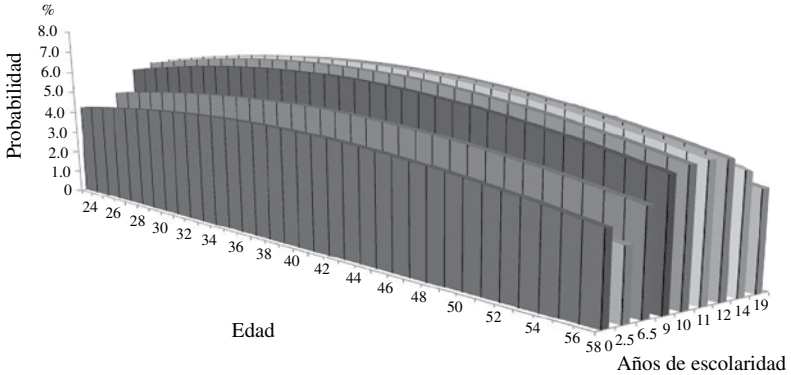
Cuadro 4  
Estimaciones con los Métodos *WESML* y *GMM*,  
hombres 24 a 58 años.<sup>a</sup> Censos de México y EU 1990

	WESML		GMM	
	Muestra Original	Muestra con Subcontados (MAE) <sup>c</sup>	Muestra Original	Muestra con Subcontados (MAE) <sup>c</sup>
N	178,078	185,599	178,078	185,599
Educación	-0.1088* (0.0063)	-0.0579* (0.0055)	-0.1088* (0.0003)	-0.0579* (0.0005)
Educación <sup>2</sup>	0.0315* (0.0016)	0.0156* (0.0014)	0.0315* (0.0001)	0.0156* (0.0001)
Educación <sup>3</sup>	-0.0025* (0.0001)	-0.0012* (0.0001)	-0.0025* (1.2E-06)	-0.0012* (0.0001)
Educación <sup>4</sup>	0.0001* (3.72E-06)	2.63E-05* (3.29E-06)	0.0001* (1.72E-06)	2.63E-05* (1.92E-06)
Edad	0.0062* (0.0022)	0.0026* (0.0020)	0.0062* (0.0002)	0.0026* (0.0002)
Edad <sup>2</sup>	-0.0001* (2.80E-05)	-0.0001* (2.52E-05)	-0.0001* (1.5E-05)	-0.0001* (1.42E-05)
Constante	2.0181* (0.0422)	2.2361* (0.0381)	-2.2485* (0.0022)	-2.1062* (0.0031)
Tasa de Muestreo en EU <sup>b</sup>	0.3301	.3478	0.3283* (.0001)	.3472* (.0001)
Tasa de Migración Promedio	.984	.987	.0532* (.0132)	.0601* (.0204)
Pseudo R <sup>2</sup> / Función objetivo	.0068	.0185	.0104	.0189
GMM Hausman**			12.23	12.34
$\chi^2$				

Fuentes: INEGI (2010), IPUMS (2010) y MMP 93.

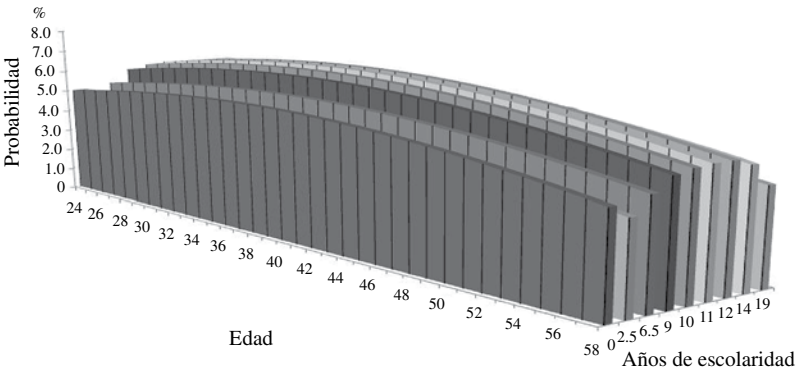
Notas: a. Todas las observaciones en la muestra de EU entraron a dicho país después de los 17 años. b. Para la estimación *WESML*, esta variable forma parte de los datos empleados en la estimación. c. Muestra aumentada usando la muestra aleatoria estratificada de acuerdo a la distribución de educación de emigrantes mexicanos ilegales, tal como se explica en la sección de los subcontados. \*Significativo al 1%. \*\*Se refiere a la prueba de no diferencia en los parámetros de los estimadores GMM y WESML para las muestras correspondientes. Se rechaza la hipótesis nula al 5%. El estadístico tiene 5 grados de libertad.

Gráfica 1  
Probabilidad de emigrar a Estados Unidos desde México.  
Hombres 24-58 años de edad, 1990. Método GMM



Fuente: Censo México 1990 y Censo EU 1990.

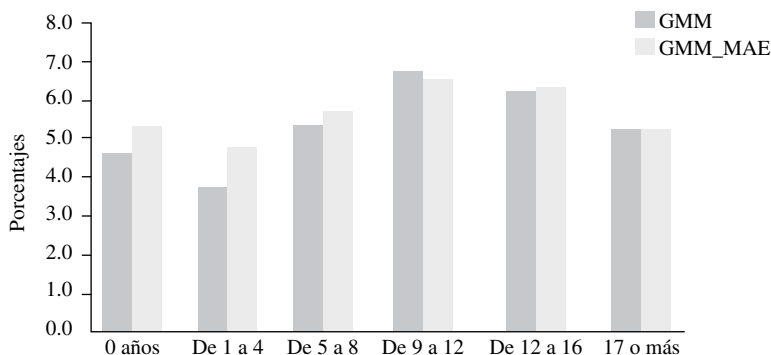
Gráfica 2  
Probabilidad de emigrar a Estados Unidos desde México.  
Hombres 24-58 años de edad, 1990. Método GMM, incluyendo  
subcontados muestra MAE



Fuente: Censo México 1990 y Censo EU 1990. Los subcontados se generan añadiendo 15% de observaciones usando muestreo estratificado de acuerdo a la distribución de educación de emigrantes ilegales obtenida del MMP 93.



Gráfica 3  
 Probabilidades de emigración promedio a Estados Unidos por nivel de educación, método GMM, muestra original y muestra MAE, México 1990



Fuente: Censo México 1990 y Censo EU 1990. Muestra MAE obtenida aumentando el número de observaciones en 15% usando muestreo estratificado de acuerdo a la distribución de educación de emigrantes ilegales a Estados Unidos, obtenida del MMP 93.

que los coeficientes *WESML* y *GMM* son estadísticamente no similares. Al igual que en la estimación para el año 1990, la diferencia radica en la constante, la cual le permite al estimador *GMM* predecir una tasa promedio de emigración similar a la observada, mientras que al estimador *WESML* queda lejos de la tasa de emigración observada. Esto hace que los estimadores *GMM* se prefieran sobre los estimadores *WESML*.

La gráfica 4 permite interpretar los resultados obtenidos de una mejor forma. Muestra que el máximo de probabilidad de emigrar ocurre a la edad de 42 años. También se muestra un polinomio en educación de cuarto grado y permite observar tres crestas ubicadas en: cero años de educación, en 11 años de educación y en 19 años de educación.<sup>20</sup>

La gráfica 5, por otro lado, muestra las tasas de emigración promedio por nivel de educación. Esta gráfica sintetiza los resultados mostrados en las gráficas 3 y 4. La gráfica sugiere que para el año 2000, las tasas de emigración del grupo de cero años y de 17 o más años de educación son los máximos, siendo la segunda el máximo global. Esto nos indica que la

<sup>20</sup> Se llevó a cabo una prueba F para saber si era necesario incluir un término de quinto grado en el polinomio. No puede rechazarse la hipótesis de que dicho elemento es cero.

Cuadro 5  
Estimaciones *WESML* y *GMM*, hombres 24 a 58 años<sup>a</sup>  
Censos de México y EU 2000

	WESML		GMM	
	Muestra Original	Muestra con Subcontados (MAE) <sup>c</sup>	Muestra Original	Muestra con Subcontados (MAE) <sup>c</sup>
N	2,176,496	2,210,947	2,176,496	2,210,947
Educación	-0.5726* (0.0114)	-0.4878* (0.0110)	-0.5726* (0.0012)	-0.4878* (0.0010)
Educación <sup>2</sup>	0.1350* (0.0027)	0.1129* (0.0027)	0.1350* (0.0021)	0.1129* (0.0002)
Educación <sup>3</sup>	-0.0107* (0.0002)	-0.0090* (0.0002)	-0.0107* (0.0001)	-0.0090* (0.0001)
Edad	0.0003* (6.20E-06)	0.0002* (6.00E-06)	0.0003* (3.40E-06)	0.0002* (3.00E-06)
Edad <sup>2</sup>	0.0780* (0.0039)	0.0684* (0.0038)	0.0780* (0.0009)	0.0684* (0.0003)
Constante	-1.4098* (4.90E-05)	-1.0875* (4.76E-05)	-2.3392* (1.10E-05)	-2.3309* (3.76E-06)
Tasa de Muestreo en EU <sup>b</sup>	.5222	.511	.5134* (.0003)	.5110* (.0006)
Tasa de Migración Promediob	.393	.439	.0924* (.0306)	.0975* (.0329)
Pseudo R <sup>2</sup> / Función objetivo GMM	.0147	.0189	.0138	.0182
Hausman** $\chi^2$			13.01	13.6

Fuentes: INEGI (2010), IPUMS (2010) y MMP 93.

<sup>a</sup>Todas las observaciones en la muestra de EU entraron a dicho país después de los 17 años.

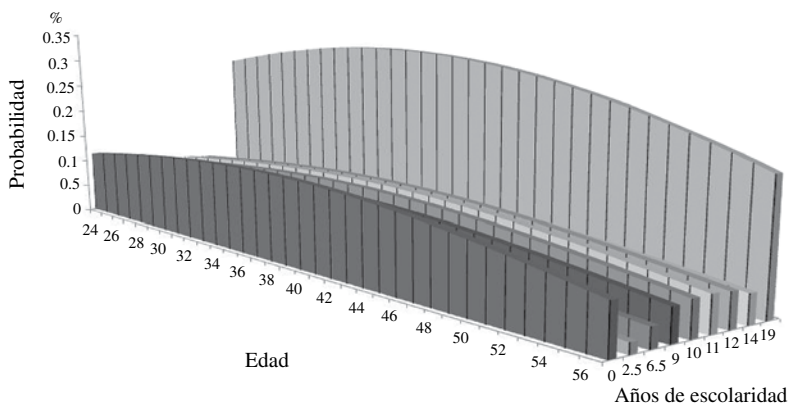
<sup>b</sup>Para la estimación *WESML*, esta variable forma parte de los datos empleados en la estimación.

<sup>c</sup>Muestra aumentada usando la muestra aleatoria estratificada de acuerdo a la distribución de educación de emigrantes mexicanos ilegales, tal como se explica en la sección de los subcontados.

\*Significativo al 1%.

\*\*Se refiere a la prueba de no diferencia en los parámetros de los estimadores GMM y *WESML* para las muestras correspondientes. Se rechaza la hipótesis nula al 1%. El estadístico tiene 5 grados de libertad.

Gráfica 4  
 Probabilidad de Emigrar a los EU desde México. Hombres 24 a 58 años de edad, 2000. Método GMM



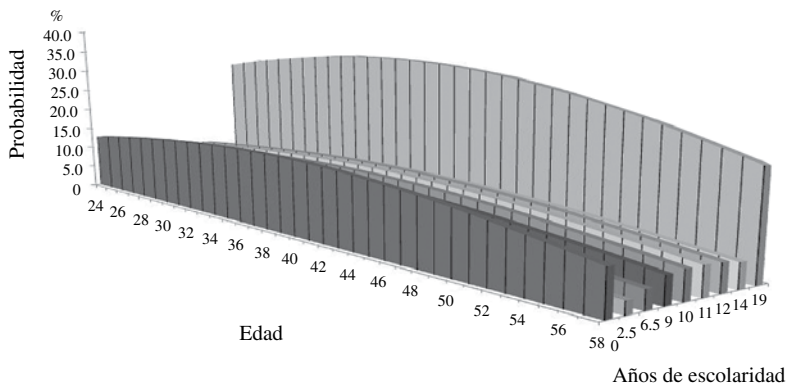
Fuente: Censo México 1990 y Censo EU 1990.

emigración cambió sustancialmente en la década de los noventa de una emigración con selección intermedia, a una emigración con selección en los extremos. La gráfica 5 también muestra que la inclusión de los subcontados mediante la muestra MAE no genera cambios cualitativos en los resultados.

#### Comparación entre censos

La columna 1 del cuadro 6 compara la probabilidad de emigrar en el año 2000 con la probabilidad de emigrar en el año 1990 por categoría educativa. El cuadro muestra cambios porcentuales. El cuadro nos muestra claramente que la emigración de mexicanos que creció más en la década de los noventa es la de los mexicanos con 17 años o más de educación, mientras que el segundo grupo que creció más en la década de los noventa es el de los mexicanos con cero años de educación. El tercer grupo con el cambio más grande es el de los mexicanos entre 9 y 12 años de educación. Esto nos muestra un cambio de selección intermedia a selección de los extremos de la distribución de educación, si bien es evidente que el grupo de mexicanos de 9 a 12 años de educación continúa creciendo en Estados Unidos. La columna 2 del cuadro 6 nos muestra el mismo cálculo realizado para la muestra MAE. Es evidente que incluir los subcontados no cambia cualitativamente este resultado.

Gráfica 5  
 Probabilidad de Emigrar a Estados Unidos desde México.  
 Hombres 24 a 58 años de edad, 2000. Método *GMM*, Muestra *MAE*



Fuente: Censo México 1990 y Censo Estados Unidos 1990. Los subcontados se generan añadiendo 10% de observaciones usando un muestreo estratificado que sigue a la distribución de educación de emergentes ilegales obtenida del MMP 93.

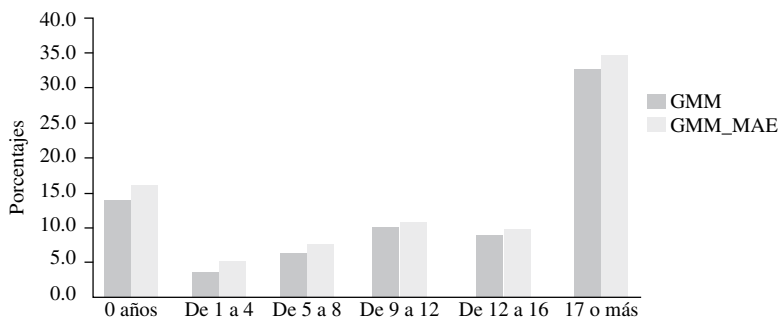
Cuadro 6  
 Cambios en la emigración México Estados Unidos 2000-1990.  
 Método *GMM*. Diferentes metodologías para corrección de subcontados y diferentes muestras originales analizadas

Años de Educación	Muestra Original	Muestra MAE	Muestra MAA	Cohorte 24 a 48 en 1990	Migrantes Recientes
0 años	198.23%	192.47%	234.01%	198.23%	161.31%
De 1 a 4	-5.66%	5.90%	7.43%	-5.66%	105.84%
De 5 a 8	15.24%	30.23%	32.81%	15.24%	78.75%
De 9 a 12	47.13%	60.93%	69.69%	47.13%	41.83%
De 12 a 16	39.20%	49.60%	60.12%	39.20%	20.19%
17 o más	516.00%	549.21%	569.75%	516.00%	295.62%

Fuentes: INEGI (2010), IPUMS (2010) y MMP 93.

Notas: Para la muestra original, la muestra MAE, la muestra MAA y la cohorte 24 a 48 años en 1990, se incluyen en la muestra de mexicanos en EU aquéllos que entraron después de los 17 años. Para la muestra de migrantes recientes sólo se incluyen aquéllos que entraron después de los 17 años y que entraron en los últimos 5 años previos al 2000. La muestra MAE se aumenta en 15% para 1990 y en 10% para el 2000, usando una muestra estratificada que sigue a la distribución de educación de ilegales que se obtiene de la muestra MMP93. La muestra MAA aumenta en 15% para 1990 y en 10% para el 2000, usando una muestra aleatoria. La Cohorte 24 a 48 en 1990 utiliza para 1990 únicamente a individuos que cumplen con este criterio de edad. En los censos del 2000 utiliza únicamente individuos entre 34 y 58 años de edad en dicho año.

Gráfica 6  
 Probabilidades de Emigración Promedio a Estados Unidos por  
 Nivel de Educación, Método *GMM*, Muestra Original y Muestra *MAE*,  
 México 2000



Fuente: Censo México 2000 y Censo Estados Unidos 2000. Muestra MAE obtenida aumentando el número de observaciones en 10% usando muestreo estratificado de acuerdo a la distribución de educación de emigrantes ilegales a los Estados Unidos, obtenida del MMP 93.

### *Análisis de robustez*

Existen tres posibles desventajas de los resultados presentados.<sup>21</sup> La primera desventaja tiene que ver con el uso de una muestra estratificada aumentada (MAE). Para analizar si la utilización de la muestra MAE sesga nuestros resultados, ahora presentamos los resultados que se obtienen usando la muestra aumentada aleatoria (MAA). La tercera columna del cuadro 6 presenta los cambios porcentuales entre censos que se obtienen al utilizar la muestra MAA. Es evidente que no hay un cambio cualitativo en los resultados.<sup>22</sup>

Una segunda desventaja tiene que ver con el hecho de que al comparar individuos entre 24 y 28 años en dos momentos del censo, en realidad no estamos haciendo una comparación sobre el mismo grupo de

<sup>21</sup> Existe una cuarta desventaja que tiene que ver con el conocido hecho de que la migración es heterogénea, pues existen algunos estados mexicanos donde la migración es mucho más intensa que en otros estados (Durand, Massey y Zenteno, 2001; Unger, 2005). Sin embargo, no existe una manera sencilla de resolver este problema, pues los datos de Estados Unidos no permiten identificar el estado mexicano de procedencia de los individuos. Se agradece a uno de los árbitros anónimos por sugerir esta potencial deficiencia en la estimación.

<sup>22</sup> Otro potencial problema que puede existir con nuestro enfoque de corregir el sub-cuento tiene que ver con la cantidad de sub-contados que se consideran. Se realizaron dos experimentos: 1) se redujo el sub-cuento a 10% en 1990 y a 5% en el 2000. 2) se aumentó el sub-cuento a 20% en 1990 y a 15% en el 2000. Se encontró que no hay cambios cualitativos en los resultados. No se incluyen en el texto por falta de espacio.

personas y, por lo tanto, reflejamos tanto cambios en la composición de la muestra como cambios en el comportamiento de los individuos. La columna 4 del cuadro 6 muestra los cálculos hechos para el grupo de personas que tenían entre 24 y 48 años de edad en 1990 y que en el censo 2000 tienen entre 34 y 58 años de edad. Es claro que no hay un cambio cualitativo en los resultados.

La tercera desventaja tiene que ver con la posibilidad de que los emigrantes mexicanos adquieran educación en Estados Unidos.<sup>23</sup> Por esta razón se eligen individuos que entraron a Estados Unidos después de los 17 años en todo nuestro análisis. Sin embargo, es posible que a esa edad todavía hayan adquirido educación en Estados Unidos, especialmente para los niveles educativos de *high school*, universidad o posgrados. Para resolver el problema presentamos un análisis en el cual, dentro de la muestra de mexicanos en Estados Unidos, dejamos a los individuos con apenas cinco años en Estados Unidos y que entraron después de los 17 años. La columna 5 del cuadro 6 nos muestra que esta vez es posible observar que el cambio en la década de los noventa es claramente decreciente en el nivel de educación, pues entre los 0 años de educación y los 16 años de educación los cambios 2000-1990 son decrecientes. Sin embargo, se obtiene nuevamente que los individuos con 17 años o más de educación presenten el crecimiento más importante de todos los grupos educativos.

Estos resultados sugieren que el comportamiento de los individuos con 17 o más años de educación es muy diferente al resto de los grupos educativos, posiblemente en consecuencia de la política migratoria de Estados Unidos que claramente los beneficia. Los resultados también sugieren que existe la presencia tanto de la emigración de retorno como la acumulación de educación de los individuos. Esto es muy claro para el grupo de educación entre 1 y 4 años de educación. Al comparar la tasa de emigración entre emigrantes recientes es la tercera más alta, pero al comparar la tasas de emigración incluyendo a todos los emigrantes es la más baja de todas. Esto nos sugiere que dicho grupo: a) retorna más frecuentemente a México o b) adquiere educación en Estados Unidos. Un análisis detallado al respecto escapa al alcance de este artículo.

#### ■ *Implicaciones de los resultados*

*Implicaciones de la autoselección educativa de los emigrantes mexicanos en EU para la política focalizada a grupos educativos específicos*

---

<sup>23</sup> Este problema potencial fue señalado por dos árbitros anónimos, a los cuales se agradece su sugerencia.

Los resultados obtenidos implican que la autoselección ha cambiado a lo largo del tiempo y que para cualquier año considerado, el tipo de selección es más complejo de lo que la teoría predice. En particular se concluye que, si bien para el año 1990 existía selección intermedia, para el año 2000 la selección se movió hacia los extremos de la distribución de educación. Sin embargo, es importante recalcar que estos movimientos para grupos educativos específicos no necesariamente revelan el impacto agregado de la emigración para Estados Unidos. Esto se hará en las implicaciones de la autoselección educativa de los emigrantes mexicanos en Estados Unidos para el agregado de las economías receptora y expulsora.

Ahora se analizará el impacto de los resultados para la política inmigratoria de Estados Unidos, focalizada a grupos específicos de educación y el impacto para México para dichos grupos. En particular, los resultados muestran una fuga de cerebros importante que para México constituye motivo de preocupación y para Estados Unidos una muestra de que su política de atracción de talentos ha funcionado.

En el caso específico de los individuos con cero años de educación representa para México una indicación de que la reducción del número de personas en México con cero años de educación no se debe al éxito de la política pública para reducir el número de personas con dichas características en México, sino a la atracción que ejerce Estados Unidos para ellos. Por otro lado, para Estados Unidos representa una consecuencia no deseada de su política inmigratoria de cerrar las fronteras, pues esto aumenta los costos de emigración, lo cual se sabe que los emigrantes responden a dicho aumento prolongando sus estadías en Estados Unidos (Reyes, Johnson y Van Swearingen, 2002; Rendon y Cuecuecha, 2010). Además se ha mostrado que el grupo con cero años se mantiene por más tiempo en Estados Unidos, lo cual genera para Estados Unidos la necesidad de pensar en políticas que le permitan a dicho grupo de emigrantes mexicanos adaptarse y asimilarse a Estados Unidos.

### *Implicaciones de la autoselección educativa de los emigrantes mexicanos en Estados Unidos para el agregado de las economías receptora y expulsora*

Como se mencionó anteriormente, los cambios encontrados en los patrones educativos de los mexicanos en Estados Unidos no necesariamente revelan el cambio que en el agregado puede darse para las economías receptoras y expulsoras. Una manera de buscar este impacto agregado es usar el enfoque de Borjas (1987), en el cual se busca detectar las zonas de la distribución de salario en las economías expulsoras y receptoras que se ven afectadas por la emigración (inmigración). Chiquiar y Hanson (2005)

plantean el uso de técnicas semi-paramétricas de estimación de salarios, para estos fines.

Usando una técnica desarrollada por DiNardo, Fortin y Lemieux (1996), identifican el área de la distribución de salarios en México que se ha perdido por la emigración. Ellos argumentan que hubo un cambio hacia la selección intermedia en la década de los noventa. En esta sección se muestra que el uso de las técnicas GMM es fundamental para estimar correctamente el cambio en selección. Específicamente se muestra que la estimación de Chiquiar y Hanson (2005) subestima el cambio en auto-selección.

Para aquéllos no familiarizados con esta literatura, se reproduce en este lugar la ecuación 9 del artículo mencionado, la cual será la ecuación 10 de este artículo:

$$(10) \quad \int \theta f^{Mex}(w/x)h(x|i = US, D_i = 1)dx$$

Donde:

$\theta f^{Mex}(w/x)h(x|i = US, D_i = 1)$ : es la densidad de la distribución de salarios en México ponderada por la densidad de las características de los emigrantes mexicanos en Estados Unidos.

$\theta = \theta^P \theta^M$ : es el ponderador que se utiliza para generar las densidades contrafactuales.

$$\theta^P = \frac{\Pr(D_i = 1|i = US, x)}{\Pr(D_i = 1|i = Mex, x)}: \text{ representa las diferencias en participación}$$

en la fuerza laboral entre individuos en Estados Unidos y México.

$$\theta^M = \frac{\Pr(i = US|x)}{\Pr(i = Mex|x)}: \text{ representa las diferencias en características entre}$$

los emigrantes y los no emigrantes.

$\theta$  también necesita tomar en consideración las tasas de muestreo de los individuos considerados en la estimación, pero los pesos se normalizan para siempre sumar 1. De acuerdo a Chiquiar y Hanson (2005), esto genera que las tasas de muestreo se eliminen siempre de la estimación y por lo tanto afirman que son irrelevantes. Chiquiar y Hanson (2005) también encontraron que la participación en la fuerza laboral no era importante en sus estimaciones y por lo tanto sólo utilizaron las diferencias en características entre emigrantes y no emigrantes.



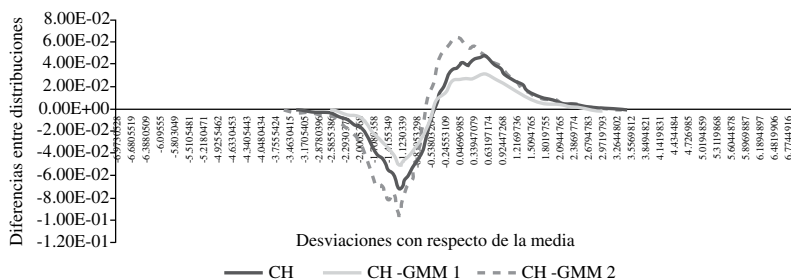
Finalmente, para generar el tipo de selección en cada año analizado se resta de la distribución no paramétrica observada para cada año, la distribución contrafactual. A continuación se muestran tres tipos de estimaciones contrafactuales que se restaron de la distribución observada:

La primera reproduce a detalle las estimaciones siguiendo las indicaciones de Chiquiar y Hanson (2005). La segunda utiliza la técnica *GMM* y utiliza los factores de expansión para cada observación. La tercera utiliza la técnica *GMM* pero no utiliza los factores de expansión para cada observación.

La gráfica 7 muestra las diferencias entre la densidad de salarios observada y las densidades contrafactuales indicadas para el año 2000. Es claro que el no usar la técnica *GMM* ni los factores de expansión genera una subestimación del tipo de selección. Sin embargo, el resultado principal de Chiquiar y Hanson (2005) de que la auto-selección para el año 2000 es intermedia permanece.

La gráfica 8 a su vez indica los mismos cálculos hechos para el año 1990. Esta vez las diferencias son imperceptibles. Esto nos sugiere que cuando la selección en características educativas es intermedia, desempeña un papel menos importante el usar las técnicas *GMM*. Pero cuando la distribución de características educativas cambia a una selección en los extremos, tenemos que el uso de la técnica *GMM* es fundamental.

Gráfica 7  
Diferencias entre distribución no paramétrica de salarios observada y distribuciones contrafactuales para ponderadores alternativos, México 2000



Fuente: Censo México 2000 y EU 2000. La línea CH utiliza la metodología de Chiquiar y Hanson (2005). CH-GMM 1 utiliza técnicas GMM y utiliza factores de identificación para cada individuo. CH-GMM 2 utiliza técnicas GMM pero no utiliza los factores de inflación para cada individuo.

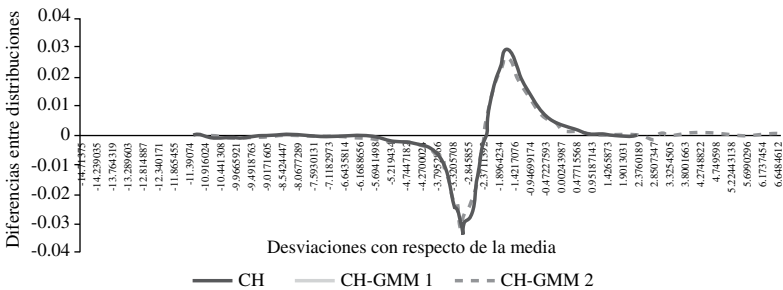
La gráfica 9 nos muestra el cambio en autoselección en la década de los noventa. Es claro que el no utilizar *GMM* genera una subestimación del cambio en autoselección. Sin embargo, el resultado de Chiquiar y Hanson (2005) de que la auto-selección cambió en la década de los noventa mostrando una mayor concentración en la parte media de los salarios mexicanos se mantiene.

Esto nos indica que, a pesar de que los cambios más importantes se dieron para los extremos de la distribución de educación, debido a que los cambios en la parte media de la distribución de educación también fueron importantes y debido a que este grupo es el más importante en términos de volumen, el cambio que domina en la auto-selección es el proveniente de la parte media de salarios.

La implicación entonces para el efecto agregado sobre Estados Unidos es que los mexicanos no generan ni una carga fiscal adicional, lo cual ocurriría si hubiera selección negativa, ni generan una ganancia en impuestos adicionales, lo cual ocurriría si hubiera selección positiva. Para el caso de México, la implicación es que la emigración muy probablemente aumenta la desigualdad salarial agregada, pues se pierde masa salarial en el centro de la distribución de salarios.

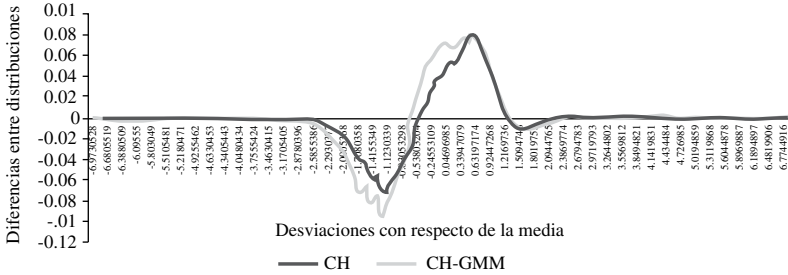
Gráfica 8

Diferencias entre distribución no paramétrica de salarios observada y distribuciones contrafactuales para ponderadores alternativos, México 1990



Fuente: Censo México 2000 y EU 1990. La línea CH utiliza la metodología de Chiquiar y Hanson (2005). CH-GMM1 utiliza técnicas GMM en la obtención de las probabilidades de migración y utiliza los factores de inflación para cada individuo. CH-GMM 2 utiliza técnicas GMM pero no utiliza los factores de inflación para cada individuo.

Gráfica 9  
Cambio en selección para el periodo 2000-1990. Métodos alternativos



Fuente: Censo México 2000 y EU 1990. La línea CH utiliza la metodología de Chiquiar y Hanson (2005). CH-GMM1 utiliza técnicas GMM en la obtención de las probabilidades de migración y utiliza los factores de inflación para cada individuo. CH-GMM 2 utiliza técnicas GMM pero no utiliza los factores de inflación para cada individuo.

*Relación de resultados con literatura existente sobre la selección de mexicanos*

Las estimaciones de este artículo son las primeras en utilizar técnicas *GMM* en datos representativos a nivel nacional. Sin embargo, existen numerosos estudios sobre las características educativas de los emigrantes mexicanos con resultados divergentes, con base en el año estudiado, la muestra estudiada, o la técnica utilizada.

Entre los que han encontrado selección negativa se encuentran los siguientes. Usando datos de encuestas rurales, Taylor (1986) encuentra que existe selección negativa hacia Estados Unidos y selección positiva hacia las ciudades mexicanas. Orrenious (1999) también encuentra selección negativa usando datos únicamente del MMP. Ibarra y Lubotsky (2007) encuentran selección negativa usando únicamente datos del censo de México 2000. Fernández-Huertas (2008) usa datos de la encuesta nacional de empleo del 2000 al 2004 y encuentra selección negativa. Kaestner y Malamud (2010) usan datos de la ENNVIH y afirman que encuentran selección negativa.

Entre los estudios que han encontrado selección intermedia se encuentran: Chiquiar y Hanson (2005), que utilizan datos del censo 2000 y las técnicas de distribuciones de salarios contrafactuales explicada en la sección anterior; Orrenius y Zavodny (2005), usando datos del MMP, encuentran selección intermedia. Ninguno de estos artículos utiliza técnicas de control de sesgo por uso de muestras endógenas. Cuecuecha

(2008) utiliza datos de la ENIGH 1994 y del CPS 1994, así como técnicas de control de sesgo por uso de muestras endógenas y encuentra selección intermedia. Sin embargo, la ENIGH 1994 tiene una muestra muy pequeña de la población en las colas de la distribución de educación, por lo que los resultados no tienen la misma validez que los resultados de este artículo que se basan en censos.

Entre los artículos que han encontrado selección positiva se encuentran: Mishra (2007), quien utiliza datos agregados de los censos de México y Estados Unidos para 1990 y 2000, encuentra selección positiva; Lowell, Passel y Pederzini (2008) utilizan datos de los censos México y Estados Unidos para el año 2000, encuentran que la tasa de emigración es creciente entre grupos educativos bajos, que la tasa de emigración es baja para individuos con licenciatura y muy alta para individuos con posgrado.

Por su parte, Cuecuecha, González y Pederzini (2010) utilizan datos de los censos 1990 y 2000, para México y Estados Unidos, así como datos del CPS, ENOE y ENE para los años 2002 a 2008. Estos autores sólo analizan la distribución de educación y sus cambios en el tiempo, así como la tasa de emigración no condicionada por grupos de edad. Ellos encuentran que la tasa de emigración no condicionada creció más para los grupos de cero años de educación y de más de 16 años de educación. Los resultados de este artículo, sin embargo, son los primeros en determinar los cambios porcentuales en las tasas de emigración condicionadas a las edades de los emigrantes, así como ajustadas por las técnicas que corrigen el sesgo generado por el uso de muestreos basados en las decisiones de los individuos.

### ■ *Conclusión*

Este artículo combina los datos de los censos de Estados Unidos y México usando técnicas econométricas que toman en consideración el hecho de que las personas eligieron el país de su residencia. Se analizan con estas técnicas la auto-selección en características educativas y el cambio en auto-selección en la década de los noventa. Este cambio en auto-selección se determina en términos de la distribución de educación, así como en términos de la distribución de salarios.

El artículo encuentra que las técnicas *GMM* son más adecuadas que las técnicas *WESML* para el estudio de la emigración de mexicanos debido a que las primeras pueden predecir mejor la tasa de emigración agregada.

El artículo también encuentra que, en 1990, la auto-selección en características educativas provenía primordialmente de individuos de la

parte media de la distribución de educación. Para el año 2000 se encuentra que la autoselección proviene de los extremos de la distribución de educación: las personas con cero años de educación y las personas con 17 o más años de educación. No sorprendentemente se muestra que el cambio a través del tiempo es máximo para estos dos grupos. Estos resultados se mantienen aun cuando se someten a diferentes análisis de robustez, incluyendo un control por el subconteo de mexicanos en los censos de Estados Unidos; un control por posibles cambios en la composición de los grupos de edad estudiados, así como por la posible adquisición de educación en este país.

Un tercer resultado de este artículo es que la aplicación de técnicas GMM para corrección de sesgo por uso de muestras endógenas es fundamental para determinar el tamaño del efecto agregado generado por la emigración. En particular se encuentra que la emigración de mexicanos a Estados Unidos proviene en el agregado de la parte media de la distribución de salarios de México.

Estos resultados informan tanto a las políticas focalizadas para grupos educativos específicos como a la política emigratoria en general. Para México, los resultados muestran que el país enfrenta un grave problema de fuga de cerebros y de que la reducción de personas con cero años de educación en México obedece a la atracción que Estados Unidos ejerce sobre ellos y no a un éxito de política pública. Para Estados Unidos representa que su política inmigratoria ha sido exitosa en atraer mexicanos con alta educación y que su política de aumentar los costos de emigración ha tenido la consecuencia no deseada de aumentar el número de mexicanos con cero años de educación en Estados Unidos, por lo cual sería recomendable que aplicara políticas que le permitan a dicho grupo asimilarse a Estados Unidos.

Finalmente, para la política emigratoria en agregado, los resultados sugieren que la emigración de mexicanos a Estados Unidos, en promedio, no representa una carga fiscal para dicho país, ni tampoco un incremento en sus ingresos fiscales. Para México, los resultados muestran que la emigración posiblemente aumenta la desigualdad salarial al reducirse la masa de mexicanos en el centro de la distribución de salarios.

## ■ *Bibliografía*

- Amemiya, Takeshi (1985). *Advanced Econometrics*. Harvard University Press. Cambridge, MA.
- Borjas, George (1987). "Self Selection and the Earnings of Immigrants". *The American Economic Review* 77(4):531-553.

- Bover, Olympia y Manuel Arellano (2002). "Learning about Migration Decisions from the Migrants: Using Complementary Datasets to Model Intra-Regional Migrations in Spain". *Journal of Population Economics* 15:357-380.
- Card, David y Ethan G. Lewis (2007). "The Difussion of Mexican Immigrants during the 1990's: Explanations and Impacts". En *Mexican Immigration to the United States*, George Borjas (editor). Chicago: The University of Chicago Press.
- Chiquiar, Daniel y Gordon H. Hanson (2005). "International Migration, Self Selection, and the Distribution of Wages". *Journal of Political Economy*. 113 (2). 239:280.
- Chiswick, Barry R. (1986). "Human Capital and the Labor Market Adjustment of Immigrants: Testing Alternative Hypothesis". En *Research in Human Capital and Development: Migration, Human Capital and Development*. Vol. 4, ed. O. Stark. Greenwich, Connecticut: JAI press, 1-26.
- Cuecuecha, Alfredo (2008). *Migration from Mexico to the US: Theory and Evidence*. Berlin: Verlag Dr. Müller.
- Cuecuecha, Alfredo; Ana González y Carla Pederzini (2010). "La Auto Selección Educativa de los Emigrantes Mexicanos a los Estados Unidos". Documento de trabajo. Universidad Iberoamericana Puebla.
- DiNardo, John, Nicole Fortin y Thomas Lemieux (1996). "Labor Market Institutions and the distribution of wages, 1973-1992: A Semiparametric Approach". *Econometrica* 64 (septiembre) 1001:1044.
- Durand, Jorge, Douglass Massey y Ricardo Zenteno (2001). "Mexican Immigration in the United States: Continuities and Changes". *Latin American Research Review* 36 (1): 107-127.
- Fernández-Huertas Moraga, J. (2008). "New Evidence on Emigrant Selection" UFAE and IAE Working Papers 742.08, Unitat de Fonaments de l'Anàlisi Econòmica (UAB) and Institut d'Anàlisi Econòmica (CSIC).
- Hausman, J. A. (1978). "Specification Tests in Econometrics". *Econometrica* 46 (6):1251-1271.
- Ibarrarán, P. y Lubotsky, D. (2007). "Mexican Immigration and Self-Selection: New Evidence from the 2000 Mexican Census". En *Mexican Immigration to the United States*, George Borjas (editor). Chicago: The University of Chicago Press.
- Imbens, Guido W. (1992). "An Efficient Method of Moments Estimator for Discrete Choice Based Models with Choice-Based Sampling". *Econometrica* 60 (5):1187-1214.
- Imbens, Guido W. y Lancaster T. (1994). "Combining Micro and Ma-

- cro Data in Microeconomic Models”. *Review of Economic Studies* 61:655-680.
- Instituto Nacional de Estadística Geografía e Informática (INEGI, 2010). Disponible en <http://www.inegi.org.mx/est/contenidos/Proyectos/ccpv/cpv2000/default.aspx>
- Integrated Public Use Microdata Series (IPUMS, 2010). Disponible en: <http://usa.ipums.org/usa/>
- Kaestner, R. y O. Malamud (2010). “Self-selection and International Migration: New Evidence from Mexico”. NBER Working Paper No. 15765.
- Lowell Lindsay y Carla Pederzini (2010). *Profesionistas Mexicanas en Estados Unidos: ¿Fuga de Cerebros?*, documento en proceso de publicación.
- Lowell Lindsay, Jeffrey Passel y Carla Pederzini (2008). “The Demography of Mexico/US Migration. En *Mexico–US Migration Management: A Binational Approach*. Agustín Escobar Latapi y Susan F. Martin (editores). Lexington Books (ISBN-13:978-0-7391-2576-2).
- Manski, Charles F. y Steve R. Lerman (1977). “The Estimation of Choice Probabilities from Choice Based Samples”. *Econometrica* 45 (8):1977-1988.
- Massey, D. S. y K. Espinosa (1997). *What’s Driving Mexico-US Migration? A Theoretical, Empirical, and Policy Analysis*. American Journal of Sociology 102 (4):939-999.
- Mckenzie, D. y H. Rapoport (2007). “Network effects and the dynamics of migration and inequality: Theory and evidence from Mexico”. *Journal of Development Economics* 84 ( 1):1-24.
- McKenzie, D. y H. Rapoport (2007a). *Self-selection patterns in Mexico-U.S. migration: The role of migration networks*. World Bank Policy Research Working Paper 4118.
- Mishra, Prachi (2007). “Effect of Emigration on Wages in Developing Countries: Evidence from Mexico”. *Journal of Development Economics* 82:180– 199.
- Orrenius, Pia M. (1999). “Family Networks, Coyote Prices and the Rural Economy in Migration from Western Mexico: 1965-1994”. Manuscrito, Federal Reserve Bank Dallas.
- Orrenius, Pia y M. Zavodny (2005). “Self Selection among undocumented immigrants from Mexico”. *Journal of Development Economics* 78 (1), 215:240.
- Passel, Jeffrey (2004). “Mexican Immigration to the US: The Latest Estimates”. *Migration Information Source*. Washington, DC: Migration Policy Institute, March.

- (2005a). *Estimates of the Size and Characteristics of the Undocumented Population*. Pew Hispanic Center Report, March 21.
- (2005b). *Unauthorized Migrants: Number and Characteristics*. Pew Hispanic Center Report, June 14.
- Passel, Jeffrey, Randy Capps y Michael Fix (2004). *Undocumented Immigrants: Facts and Figures*. Urban Institute Immigration Studies Program Fact Sheet, January 12.
- Passel, Jeffrey S. y Michael F. Fix (2001). “US Immigration at the Beginning of the 21st Century”. Testimony Before the Subcommittee on Immigration and Claims Hearing on “The US Population and Immigration”. Committee on the Judiciary, US House of Representatives, Urban Institute, Washington, DC.
- Passel, Jeffrey, Jennifer Van Hook y Frank Bean (2004). “Estimates of the Legal and Unauthorized Foreign Born Population for the United States and Selected States, Based on Census 2000”.
- Rendon, Silvio y Alfredo Cuecuecha (2010). “International Job Search: Mexicans in and out of the U.S.” *Journal of Household Economics* 8 (1):53-82.
- Reyes, Belinda I., Johnson, H.P. y Richard Van Swaeringen (2002). *Holding the Line? The Effect of the Recent Border Build-up on Unauthorized Migration*. San Francisco, CA: Public Policy Institute of California.
- Taylor, J. Edward (1986). “Differential Migration, Networks, Information and Risk”. En *Research in Human Capital and Development: Migration, Human Capital and Development*. Vol. 4, ed. O. Stark. Greenwich, Connecticut: JAI press, 147-171.
- Unger, Kurt (2005). *Regional Economic Development and Mexican Out-Migration*. NBER Working Paper 11432.



## ■ Apéndices

### *Los datos de educación en los censos de México y EU*

Los datos de educación de los censos de ambos países no son totalmente compatibles. Seguimos las reglas de homogenización que Cuecuecha, González y Pederzini (2010) utilizan en su apéndice 1. Esto equivale a agrupar los niveles educativos en las siguientes categorías: 0) sin educación, 1) entre 1 y 4 años de educación, 2) entre 5 y 8 años de educación, 3) 9 años, 4) 10 años, 5) 11 años, 6) 12 años, 7) entre 13 y 16 años de educación, 8) más de 16 años de educación.

El cuadro A1 muestra las fracciones de la población que corresponden a cada grupo de mexicanos y para cada año con el que se cuenta. Un análisis detallado sobre la distribución de educación de mexicanos en México y en Estados Unidos se encuentra en Cuecuecha, González y Pederzini (2010).

Cuadro A1  
Distribución de educación

Años de educación	1990		2000	
	México	EU	México	EU
0	13.66	12.65	5.8	10.27
Entre 1 y 4	23.48	14.98	16.53	6.35
Entre 5 y 8	28.09	31.75	27.18	28.73
9	11.36	6.77	18.56	8.78
10	1.54	2.93	1.85	3.25
11	2.85	2.34	2.39	3.21
12	6.66	16.58	10.78	23.57
Entre 13 y 16	10.95	10.53	16.59	13.92
17 o más	1.41	1.47	0.32	1.91

Fuente: Censos México 1990 y 2000, Censos Estados Unidos 1990 y 2000.

### *El año de entrada a Estados Unidos*

Para calcular la edad de entrada a Estados Unidos se utiliza el año de entrada a Estados Unidos. Con este dato podemos separar a los individuos que entraron antes de los 17 años, pues se sabe que ese grupo de mexicanos tiene mayor probabilidad de adquirir educación en este país (Lowell y Pederzini, 2010). El censo registra el año de entrada de las personas y en otros casos registra los años que las personas tienen en Estados Unidos. La segunda variable es categórica y abarca periodos de

años. En estos casos estimamos los años en Estados Unidos iguales al punto medio de años en el intervalo considerado.

### *La distribución de educación para los mexicanos en el MMP93*

El cuadro A2 muestra la distribución de educación para los mexicanos que se encuentran en la encuesta MMP 93. La muestra incluye sólo hombres entre 24 y 58 años de edad que entraron a EU después de los 17 años. Se considera ilegales a todos los individuos que declaran en la encuesta haber cruzado a Estados Unidos sin un permiso de trabajo. El cuadro muestra al grupo de personas tomado de la muestra de 1982 a 1992 y al grupo de personas tomado de las encuestas de 1992 a 2002.

**Cuadro A2**  
Distribución de educación para mexicanos hombres  
entre 24 y 58 años de edad\*, MMP93

Años de la encuesta	No Emigrantes		Emigrantes			
	1982-1992	1992-2002	1982-1992		1992-2002	
Años de educación			To-	Ilegales	Todos	Ilegales
			dos			
0 años	14.33	7.71	21.73	20.83	11.43	11.41
1 a 4 años	31.46	18.90	41.64	42.53	34.05	32.54
5 a 8 años	28.76	31.85	25.50	25.93	33.27	33.63
9 a 12 años	14.68	26.34	7.89	6.97	16.20	17.02
12 a 16 años	7.62	8.16	2.59	3.05	3.21	3.14
17 o más	3.16	7.04	.65	.69	1.84	2.26
N	1997	4457	1698	1018	2179	1463

Fuente: MMP 93.

\*La edad se refiere para los no migrantes al momento de la encuesta; para los emigrantes se refiere a la edad que tenían en su último viaje. Sólo se incluyen individuos que emigraron después de los 17 años.