

# Modelando la migración interestatal de México: cadenas de Markov estáticas *versus* Cadenas de Markov dinámicas con medias móviles

Carlos GARROCHO, Eduardo JIMÉNEZ-LÓPEZ  
y José Antonio ÁLVAREZ-LOBATO

*El Colegio Mexiquense, México*

## *Resumen*

Se compara la precisión / imprecisión de las Cadenas de Markov Estáticas (CME) *versus* las Cadenas de Markov Dinámicas con Medias Móviles (CMD) para reproducir los flujos migratorios interestatales de México. En México aún no se ha realizado una comparación detallada de ambos modelos. La principal aportación de este texto es la manera sistemática como se compara la precisión de los dos modelos markovianos. Se detallan las métricas de bondad de ajuste para comparar la exactitud / inexactitud de las CME y las CMD. Cada métrica se orienta a una *dimensión distinta* del sistema migratorio: i. La *Desigualdad migratoria* se analiza con índices de Gini y Curvas de Lorenz; ii. La *Competencia migratoria* con índices de Herfindahl-Hirschman; iii. La *Similitud Estadística* con Coeficientes de Correlación de Pearson y de Correlación de Rangos de Spearman; y, iv. El *Error Global* con el *error proporcional promedio*. El modelo derivado de las CMD resultó ligeramente más preciso que el modelo estático para replicar los flujos migratorios interestatales observados de México y seguramente será aún más exacto conforme se disponga de más información y se amplie el periodo de estudio.

*Palabras clave:* Migración, México, Cadenas de Markov Estáticas *versus* Cadenas de Markov Dinámicas con Medias Móviles.

## *Abstract*

*Modelling interstate migration in México: Static Markov Chains versus Dynamic Markov Chains with Moving Average*

This paper compares the accuracy / inaccuracy of Static Markov Chains (CME) *versus* Dynamic Markov Chains with Moving Averages (CMD) to reproduce interstate migration flows in Mexico. This comparison has not been done in Mexico. The main contribution of this paper is the systematic way the precision of the two Markov models is compared. We detailed the metrics of comparison. Each metric is geared to a different dimension of the migration system: i. *Inequality of immigration* is analyzed with Gini indexes and Lorenz curves; ii. *Competition of migration* is compared with the Herfindahl-Hirschman index; iii. *Statistical Similarity* is examined with Pearson correlation coefficients and Spearman rank correlation; and iv. The *Global Error* is estimated with the proportional average error index. The model derived from the CMD was slightly more accurate than the static model to replicate interstate migration flows observed in Mexico and will surely be even more accurate as more information becomes available and the study period is extended.

*Key words:* Migration, Mexico, Static Markov Chains *versus* Dynamic Markov Chains with Moving Averages.

## INTRODUCCIÓN

Uno de los problemas más complicados que enfrenta la geodemografía contemporánea es develar y anticipar la distribución espaciotemporal de la población (Garrocho, 2011). La razón es que de los cuatro componentes básicos del cambio demográfico: nacimientos, defunciones, inmigraciones y emigraciones, los dos primeros han perdido importancia *relativa* como factores explicativos y son más factibles de anticipar, mientras que los dos últimos se han vuelto más relevantes, volátiles y difíciles de prever (Goodkind y West, 2002).<sup>1</sup>

El tema es relevante, ya que conocer la distribución de *la demanda* en el territorio y en el tiempo es clave para la toma de decisiones de los gobiernos y de los diversos agentes económicos. Fallas de información en este tema generan incertidumbre y limitan la planeación estratégica y táctica tanto del sector público como del sector privado, lo que al final produce ineficacia gubernamental e ineficiencia social y económica. Si bien el mercado emite señales instantáneas y reacciona rápido, los gobiernos no (Geys, 2006), y pueden actuar erráticamente durante mucho tiempo, localizando y dimensionando equivocadamente sus inversiones en términos *espaciales* (i.e. localizando servicios en lugares erróneos), *temporales* (i.e. abriendo o cerrando servicios en momentos equivocados), *sectoriales* (i.e. asignando recursos en sectores incorrectos y poco interrelacionados) o de las características del *mercado consumidor* (i.e. estimando erróneamente sus dimensiones y otras peculiaridades de la demanda que definen los segmentos del mercado). En consecuencia, resulta fundamental, tanto para el sector público como para el privado, no sólo tener una *imagen correcta* de la distribución espaciotemporal de la población, sino *anticipar* sus patrones de distribución para tomar decisiones más acertadas, asertivas y oportunas (Harvir *et al.*, 2005).

Para avanzar en la solución del enigma de la distribución espaciotemporal de la población, una clave es descubrir los patrones de migración que generan cambios notables de la demanda por servicios y oportunidades de desarrollo, tanto en el espacio como en el tiempo (Bell *et al.*, 2015).<sup>2</sup> Revi-

<sup>1</sup> Un elemento adicional: los movimientos migratorios afectan directamente los patrones espaciales y temporales de los nacimientos y las defunciones (Bustos, 2013).

<sup>2</sup> La demanda espaciotemporal por servicios y oportunidades de desarrollo también es influida por la movilidad de la población. Véase una detallada discusión sobre migración, movilidad y población flotante en Garrocho (2011).

sando la literatura internacional, es posible establecer cuatro categorías de modelos que permiten replicar los flujos migratorios (elemento crucial de la distribución espaciotemporal de la población): i) Modelos de Markov; ii) Modelos de interacción espacial; iii) Modelos econométricos; y, iv) Modelos fundamentados en la llamada ecuación master (Martínez-Faura y Gómez-García, 2001). Este trabajo se enfoca en los Modelos de Markov. Específicamente, a comparar la precisión (o imprecisión) de las Cadenas de Markov Estáticas *versus* las Cadenas de Markov Dinámicas con Medias Móviles, para reproducir los flujos migratorios interestatales de México.

Los modelos *markovianos* se fundamentan en los procesos de Markov y son de naturaleza *sistémica* (i.e. consideran simultáneamente todos los orígenes y destinos de los flujos migratorios) y básicamente *predictiva*, aunque aportan poco en términos *explicativos*. Proyectan los saldos migratorios de *todos* los orígenes y destinos, en un entorno en el que *interactúan* de manera *simultánea*. Esto implica que se compensan exactamente las entradas y salidas de migrantes de cada unidad espacial que se utilice (e.g. entidades federativas, como en este trabajo), lo que usualmente genera proyecciones de flujos migratorios con precisión aceptable (para México ver, por ejemplo, Bustos, 2013).

## OBJETIVO Y ESTRATEGIA DE PRESENTACIÓN

El objetivo de este artículo es *metodológico*: consiste en comparar la precisión (o imprecisión) de las Cadenas de Markov Estáticas *versus* las Cadenas de Markov Dinámicas con Medias Móviles, para reproducir los flujos migratorios interestatales de México. No se pretende, de ningún modo, proponer correcciones a los métodos markoviano utilizados en este texto ni, mucho menos, proponer un método nuevo de modelización y proyección de flujos migratorios. Nuestra pregunta de investigación es mucho más modesta: ¿Cuál método markoviano de modelación de flujos migratorios interestatales es más preciso (o impreciso) para México: las Cadenas de Markov Estáticas o las Cadenas de Markov Dinámicas con Medias Móviles? La pregunta, aunque sencilla, es relevante en términos metodológicos, porque permite orientar los esfuerzos analíticos de la migración en México que se apoyen en los razonamientos de las Cadenas de Markov. Si bien la literatura reporta que las CMD modela mejor los sistemas migratorios que las CME (Guijarro y Hierro, 2005), aún no se ha realizado una comparación sistemática de ambos modelos para México. Quizá la principal aportación metodológica de este texto es la manera sistemática como se compara la precisión de los dos modelos markovianos.

En lo que sigue, el artículo se estructura en cinco secciones. En la primera se explican los métodos de estimación de las Cadenas de Markov Estáticas (CME) y de las Cadenas de Markov Dinámicas con Medias Móviles (CMD). Se incluyen sus principales supuestos, ventajas y debilidades, y se sugiere cómo se pueden contrarrestar algunos de sus inconvenientes. En la segunda sección se detallan las métricas de bondad de ajuste que se utilizan en este trabajo para comparar la exactitud / inexactitud de las CME y las CMD. Utilizar varias métricas de bondad de ajuste puede generar “la paradoja de los relojes desincronizados”: ¿cuál marca la hora correcta? (ver por ejemplo los artículos clásicos de Plane y Mulligan, 1997 y el de Rogers y Sweeney, 1998). Para evitarla, cada métrica que utilizamos se orienta a una *dimensión distinta* del sistema migratorio: i) la *Desigualdad* migratoria la analizamos con índices de Gini y Curvas de Lorenz; ii) la *Competencia migratoria* con índices de Herfindahl-Hirschman; iii) la *Similitud Estadística* con Coeficientes de Correlación de Pearson y de Correlación de Rangos de Spearman; y iv) el *Error Global* con el *error proporcional promedio*.<sup>3</sup> En la tercera sección se puntualizan las fuentes de información (incluyendo un comentario sobre sus limitaciones) y el *software* utilizado. Estas tres primeras secciones son clave por el carácter metodológico del artículo. En la sección cuatro enfrentamos sistemáticamente las CME *versus* las CMD, y comparamos la exactitud / inexactitud con que reproducen el sistema migratorio mexicano. Finalmente, en la sección cinco se sintetizan los hallazgos, las aportaciones y las limitaciones y ventajas de los métodos analizados.

## MIGRACIÓN Y LOS MODELOS DE MARKOV

El método markoviano usualmente supone *homogeneidad temporal* del fenómeno que intenta modelar, por lo que estima las probabilidades de transición (i.e. de *cambio* de una situación a otra: o de *moverse* de un lugar a otro) con técnicas de tipo estático (Modica y Poggiolini, 2013). Sin

<sup>3</sup> Pare evitar el “*Problema de los relojes desincronizados*” en este texto no estimamos el Coeficiente de Variación (CV) ni el Índice de Hoover, que se traslapan con el Índice de Gini (IG) (Mulligan, 1991; Plane y Mulligan, 1997). Rogers y Sweeney (1998) entablaron una polémica con Plane y Mulligan (1997) sobre cuál de los dos indicadores era mejor para medir la distribución espacial de la migración. El tiempo le dio, en parte, la razón a Plane y Mulligan: actualmente el IG es mucho más popular que el CV. Adicionalmente, López-Vega y Velarde (2011: 130) reportan: “Rogers y Sweeney (1998: 23) consideran que el CV es un indicador mucho más transparente en términos de su lectura e interpretación. Sin embargo, en el caso de la migración interna en México, ambos (el CV y el IG) son sensibles a las características del cambio en los flujos de la migración reciente...” (El texto entre corchetes es nuestro). Todo esto sugiere que la selección de los indicadores para explorar la concentración / dispersión de los flujos migratorios depende de los objetivos de cada investigación.

embargo, la migración es un fenómeno socioespacial altamente dinámico que está influido por una serie de factores cuyo efecto no siempre es bien conocido y resulta muy complicado de anticipar (Castles *et al.*, 2013). Quizá por estas razones, las Cadenas de Markov, utilizadas desde los años ochenta en México (ver el excelente aporte pionero de Partida, 1989), han perdido cierta fuerza en años recientes como instrumento de *modelación y proyección de flujos migratorios*, pero no como instrumentos de *proyección de población* (en este tema resulta clave revisar los notables aportes que hace Bustos, 2013 al caso mexicano).

Sin embargo, recientemente se ha propuesto una mejora a las CME (que se apoyan en la hipótesis de homogeneidad temporal), mediante un *método dinámico de Medias Móviles*, que, se arguye, modela mejor los sistemas migratorios que el método estático habitual de tendencia lineal (Guijarro y Hierro, 2005). Hasta donde sabemos, las Cadenas de Markov con Medias Móviles no se han aplicado a sistemas migratorios en México, ni en América Latina, por lo que aún no han sido evaluadas sistemáticamente en la región sus presuntas ventajas respecto a las CME.<sup>4</sup>

La aplicación de métodos de Medias Móviles a las Cadenas de Markov están revitalizando su uso para modelar y proyectar sistemas migratorios (Guijarro y Hierro, 2005; Hierro y Guijarro, 2006), quizá porque están demostrando su poder predictivo en diversas áreas del conocimiento: economía (Startz, 2008), finanzas y análisis bursátiles (Cont y Wagalath, 2013; Kayacan *et al.*, 2010), genética, biología o medicina (Lin *et al.*, 2012), por mencionar algunos ejemplos.

### Las Cadenas de Markov

En este trabajo entendemos por Cadenas de Markov Estáticas las que se apoyan en la hipótesis de homogeneidad temporal. Su elemento clave son las *probabilidades de transición*: la *probabilidad condicionada de transitar* (i.e. *migrar*) de una unidad espacial a otra (i.e. estados o municipios, por ejemplo), en *dos puntos en el tiempo*.

La literatura reporta tres técnicas principales para estimar Cadenas de Markov homogéneas en el tiempo: el procedimiento *markoviano estático*, el método de *matriz homogénea* y el método de *matriz promedio*. Los tres procedimientos generan resultados similares, ya que los tres estiman una matriz con probabilidades de transición constantes, a partir de la matriz de transición teórica que caracteriza la cadena. Luego, la matriz de transi-

<sup>4</sup> En la literatura internacional el término *medias móviles* se reporta como *moving average*.

ción constante *calculada* se utiliza para *proyectar* matrices de transición en puntos consecutivos en el tiempo (Guijarro y Hierro, 2005).

Puede adelantarse, como hipótesis, que las Cadenas de Markov Estáticas (i.e. homogéneas en el tiempo) quizá no sean la mejor alternativa para modelar sistemas migratorios tan dinámicos como el mexicano, justamente por su *homogeneidad temporal* (Romo *et al.*, 2012). La alternativa, entonces, sería optar por una perspectiva *dinámica* que estime las probabilidades de transición mediante una relación *vinculada al tiempo*. Quizá la opción más conocida es la de tendencia lineal que tiene sus antecedentes en Rogerson (1979) y que ha sido aplicada con cierto éxito en realidades relativamente cercanas a la mexicana (Gómez-García *et al.*, 1997; Hierro y Guijarro, 2007). Si esto es así, se puede extender la hipótesis y plantear que es muy probable que el método de Medias Móviles sea más preciso que el estático markoviano para modelar el *sistema migratorio* de México.<sup>5</sup>

La ventaja del método de Medias Móviles aplicado a la modelización de la migración es que genera una estimación suavizada de los flujos migratorios, mediante el cálculo repetido de valores medios. Como el propósito es *suavizar la tendencia* de los flujos migratorios, matizando los cambios bruscos pero manteniendo las tendencias, el proceso consiste en calcular consecutiva e iterativamente medias aritméticas de datos sucesivos en el tiempo. El número de datos sucesivos es el llamado *orden* de la media aritmética: al final se selecciona el orden que *sea más exacto según las medidas de bondad de ajuste* utilizadas y se van *sustituyendo* los valores de la matriz de flujos calculados por los valores de las medias aritméticas que se van estimando (*valores suavizados*: Guijarro y Hierro 2005).

La selección del orden de la media móvil (i.e. la dimensión de la *ventana temporal* de observación), no es asunto sencillo (Vandewalle *et al.*, 1999). Cuando el orden es muy bajo, usualmente la media móvil se vincula fuertemente a los cambios en las probabilidades de transición, lo que le permite al método hacer ajustes frente a cambios dinámico bruscos, pero registra el inconveniente de que puede prender *alarmas falsas* ante cambios dinámicos pequeños, que se podrían ajustar en momentos subsiguientes. En el sentido opuesto, un orden elevado de la media móvil es adecuado cuando la dinámica de la cadena (v.g. *la tendencia*) es clara, ya que las correcciones que realice la media móvil serán menores (Alessio *et al.*, 2002). Sin embargo, en esta situación se sacrifica una reacción más temprana del

<sup>5</sup> Ahora bien, si se trata de *proyectar población* en México mediante Cadenas de Markov, una de las mejores alternativas (si no es que la mejor hasta el momento) sería apoyarse en el método de Bustos (2013) que permite realizar proyecciones por edad, sexo y región, incorporando explícitamente las dinámicas demográficas de orígenes y destinos.

método ante un cambio dinámico brusco (Chakraborti y Germano, 2010). De aquí la importancia de las métricas que utilizamos en este trabajo para estimar la *bondad de ajuste* de los flujos calculados respecto a los observados (ver Introducción).

### Cadenas de Markov Estáticas: definición y forma de cálculo

Las Cadenas de Markov representan uno de los procesos estocásticos más útiles para modelar fenómenos a partir de una evolución probabilística en el futuro, conociendo solamente la situación presente. Un *proceso estocástico* registra un comportamiento *no-determinista*. Es decir, la trayectoria o evolución del proceso depende tanto de variables *inherentes o causales* del proceso, como de variables *aleatorias o estocásticas*. En matemáticas, un proceso estocástico es un concepto útil para develar el *comportamiento* de variables aleatorias que cambian (evolucionan) en función del *tiempo* (aunque en ocasiones podría tratarse de otra variable: el espacio, por ejemplo) (Ching *et al.*, 2013).

Los flujos migratorios (y muchos otros procesos) observados en el tiempo son a menudo modelados mediante procesos estocásticos, entendidos como cualquier colección de variables aleatorias  $\{X(t)\}$  dependientes del tiempo  $t$  (Cameron y Poot, 2011). El tiempo puede ser discreto, por ejemplo  $t = 0, 1, 2, \dots$ , o continuo:  $t \geq 0$ . En cualquier momento, describe la observación de una variable aleatoria que denotaremos  $X_t$  o  $X(t)$ . Sea  $\{X_n\}_{n \geq 0}$  un proceso estocástico discreto con espacio de estados contable  $E = \{i, j, k, \dots\}$ . Si para todos los enteros  $n \geq 0$  y todos los estados  $i_0, i_1, \dots, i_{n-1}, i, j$ , entonces:

$$P(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) = P(X_{n+1} = j | X_n = i) \quad (1)$$

Donde ambos lados de la ecuación están bien definidos. Entonces este proceso estocástico es llamado una Cadena de Markov o un proceso de Markov, y se le llama cadena homogénea si el lado derecho de la ecuación (1) es independiente de  $n$ . La ecuación (2) son las *propiedades de Markov* para todos los estados  $i, j$  (Yin y Zhang, 2010).

$$p_{ij} \geq 0, \sum_{j \in E} p_{ij} = 1 \quad (2)$$

El espacio de situaciones puede ser infinito, y por lo tanto la matriz no es en general estudiada mediante álgebra lineal. Sin embargo, satisfacen

las operaciones básicas de suma y multiplicación. Así,  $p_i = (p_{ij})_{j \in E}$  es un vector de probabilidades para cada  $i$ .

### Cadenas de Markov Dinámicas de Medias Móviles: definición y forma de cálculo

Mientras las CME intentan modelar fenómenos a partir de la situación presente, el método de las CMD abre un “ventana temporal de observación” y considera la información *presente y la pasada*. Es decir, considera *la trayectoria histórica del proceso*. Esto le permite al Método de Medias Móviles obtener una matriz de transición teórica que caracteriza la cadena, así como una proyección de la matriz de transición para instantes consecutivos en el tiempo. Para este artículo, dada una secuencia temporal de flujos migratorios, la estimación de la correspondiente probabilidad de transición es la media aritmética de los flujos migratorios entre los distintos puntos consecutivos en el tiempo que constituyen el periodo de observación (Guijarro y Hierro, 2005).

Cuando se usa el método de Medias Móviles se está suponiendo que todas las observaciones de la serie de tiempo son igualmente importantes para la estimación del parámetro a modelar y/o pronosticar (en este caso el parámetro es la media aritmética de los flujos, según el *orden (n)* seleccionado). De esta manera, se utiliza el promedio de los  $n$  valores considerados en la “ventana de observación temporal” para encontrar la *tendencia* o *trayectoria* de los datos de migración observados.

En términos matemáticos la Media Móvil se estima de la siguiente forma:

$$\text{Media Móvil} = \frac{\sum(\text{datos en "n" puntos en el tiempo})}{\text{"n"}} \quad (3)$$

El término *móvil* indica que conforme se tiene una nueva observación de la serie de tiempo (una *nueva ventana temporal*), ésta reemplaza la observación más antigua en la ecuación y se calcula una nueva media aritmética. Dada una proyección de la matriz de flujos migratorios con la ayuda de la ecuación (1), se realizó un procedimiento de medias móviles de orden  $n$  ( $n = 2, 3, \dots, k$ ) para todo,  $i, j \in E$ , como estimación de la probabilidad de transición entre los instantes  $t$  y  $t + 1$ . Esto implica que se prueban múltiples valores de  $n$ :

$$\overline{P}_{ij}(t, t + 1) = \frac{1}{n} \sum_{u=1}^n P_{ij} \quad (4)$$



La utilización de la técnica de Medias Móviles ayuda a encontrar y suavizar la tendencia de las probabilidades en la matriz de transición, de acuerdo con su trayectoria o comportamiento histórico, lo que permite filtrar variaciones no significativas en los flujos migratorios. Esto constituye una ventaja sobre las CME que sólo consideran el presente.

### **Cadenas de Markov Estáticas: ventajas, limitaciones y soluciones**

Es importante aclarar las principales *ventajas y desventajas* de las Cadenas de Markov, para apreciar los resultados de este trabajo en su justa dimensión. La utilización de las Cadenas de Markov para modelar flujos migratorios se fundamenta en las siguientes hipótesis, algunas de las cuales constituyen sus principales *desventajas*:

- i. La formulación de las Cadenas de Markov suponen que la *población es homogénea*. Esto implica que en estudios agregados, en realidad se modelan poblaciones *heterogéneas* como si fueran *homogéneas*. Al considerar una sola matriz de transición y una sola tasa de movilidad, la homogeneidad de la población que consideran las Cadenas de Markov no reflejan completamente la realidad (esto es común a todos los modelos: lo único que refleja completamente la realidad es la realidad misma). Usualmente las Cadenas de Markov no consideran la descomposición de la población en grupos que se distinguen por edad, sexo, ingreso, entre otras variables relevantes (Rabiner y Juang, 1986; Rabiner, 1989; Giménez *et al.*, 2012). Quizá la mejor alternativa para resolver este punto para el caso de México es seguir la estrategia de Bustos, 2013.
- ii. La hipótesis de *homogeneidad temporal* está condicionada a la garantía de que la cadena es discreta, finita y regular, y que posee una distribución en equilibrio (Martínez y Gómez 2001). Una estrategia para eludir la homogeneidad temporal es, justamente, aplicar el método de medias móviles (Le Gallo y Dall'erba, 2006).
- iii. Los cambios de estado (o situación) ocurren en un tiempo discreto, es decir, las transiciones tienen lugar en intervalos regulares de tiempo. Sin embargo, en algunos casos los cambios entre estados (o situaciones) no ocurren regularmente, sino que puede ser el resultado de otras observaciones que tienen su propia distribución de probabilidad. El método de Medias Móviles ayuda a reducir este problema porque considera las tendencias a lo largo del tiempo.

Las Cadenas de Markov también tienen *ventajas* para analizar sistemas migratorios, algunas importantes, por ejemplo:

- a) Tienen un enfoque sistémico que permite considerar todos los orígenes, destinos y flujos de manera simultánea, así como la dependencia entre los flujos. Por tanto, las Cadenas de Markov son capaces de ofrecer una visión de los cambios más probables de sistemas muy complejos, como el migratorio, con una visión longitudinal (Guijarro y Hierro, 2007).
- b) Su forma de cálculo no es extremadamente compleja y menos ahora que existe *software* diverso: Java Modelling Tools (JMT),<sup>6</sup> MARCA (MARKov Chain Analyzer),<sup>7</sup> o MARCH,<sup>8</sup> e incluso paquetes de cómputo más convencionales como MatLab,<sup>9</sup> entre muchos otros.
- c) El enfoque Markoviano tiene capacidad predictiva sin necesidad de conocer las *causas* de los fenómenos que modela (los considera como procesos estocásticos). En este sentido son un instrumento valioso capaz de dar luz sobre la migración futura en el corto plazo (cinco o diez años).

## MÉTRICAS DE BONDAD DE AJUSTE

A pesar de las frecuentes referencias en la literatura demográfica a la distribución espacial de los flujos migratorios (*spatial focus*: el grado de concentración de la migración entre unidades espaciales de origen y de destino: O’Kelly y Horner, 2003), poco se ha avanzado al respecto. Estudios contemporáneos (e.g. López-Vega y Velarde, 2011) siguen utilizando los mismos indicadores que ya reportaban Plane y Mulligan (1997) y Rogers y Sweeney (1998) desde hace casi dos décadas (De Castro, 2007).<sup>10</sup> Por eso aquí introducimos también el índice de Herfindahl-Hirschman, que ha demostrado su potencia en la literatura económica sobre competencia de mercados (en el caso de este trabajo: *competencia migratoria* entre entidades federativas) y que complementa bien al índice de Gini (D’Amico y Di Biase, 2010). También utilizamos la Curva de Lorenz (CL), ya que el índice de Gini puede generar el mismo valor para distribuciones distintas (v.g. para paisajes migratorios diferentes: Giorgi, 1993). En lo que sigue se

<sup>6</sup> <https://sites.google.com/site/jaracohe/softwareparaestudio>

<sup>7</sup> <http://www4.ncsu.edu/~billy/MARCA/marca.html>

<sup>8</sup> <http://www.andreberchtold.com/march.html>

<sup>9</sup> <http://www.mathworks.com/products/matlab/>

<sup>10</sup> En realidad algunos de estos indicadores datan de la década de los años ochenta del siglo pasado (Watkins, 1986), como bien lo reportan Plane y Mulligan (1997).

detallan los indicadores de bondad de ajuste utilizados en este trabajo y el razonamiento que justifica su utilización.

### Índice de Gini y Curva de Lorenz

En este texto el índice de Gini ( $I_G$ ) es una medida de *desigualdad* de la distribución de los flujos de inmigrantes y emigrantes entre las entidades federativas y se estima de la siguiente forma:

$$I_G = \frac{\sum_{i=1}^N (p_i - q_i)}{\sum_{i=1}^N p_i} \quad (5)$$

Donde  $p_i$  y  $q_i$  representan frecuencias de ocurrencia de dos variables en un conjunto de datos (1, 2, ...,  $i$ , ...,  $N$ ). El índice de Gini no puede ser negativo y oscila entre 0 y 1. Si  $p_i = q_i$ , significa que la variable analizada se distribuye por igual entre todos los elementos (implica *máxima igualdad* entre los elementos:  $I_G = 0$ ). Por su parte,  $I_G = 1$  representa *desigualdad máxima* entre todos los elementos  $p_i \neq q_i$  (Raskall y Matehson, 1992; Spicker *et al.*, 2007).<sup>11</sup>

Adicionalmente al Índice de Gini tradicional, en este trabajo seguimos el ejemplo de López-Vega y Velarde (2011) y estimamos el índice de Gini *Global* con la metodología de Plane y Mulligan (1997), que junto con el *error proporcional promedio* son los indicadores globales utilizados en este trabajo (ver más adelante la explicación del error proporcional promedio).<sup>12</sup>

El índice de Gini se vincula directamente con la Curva de Lorenz (CL) que es una representación gráfica cartesiana de la desigualdad. En este trabajo la CL se bosqueja situando en el eje  $X$  los valores acumulados de los *migrantes observados* expresados en porcentaje respecto al total de migrantes y en el eje  $Y$  los valores acumulados de los *migrantes calculados* expresados también en porcentajes respecto al total. Si ambas series de valores acumulados son exactamente iguales no se produce una curva (la CL) sino una línea recta diagonal que va de la esquina inferior izquierda

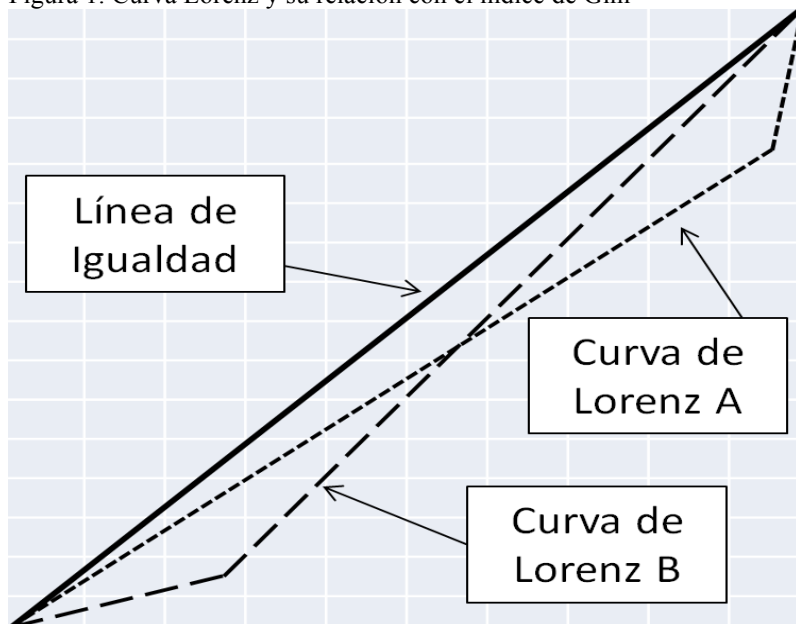
<sup>11</sup> En el excelente trabajo de López-Vega y Velarde (2011; p. 130) se registra un error en su explicación cuando dicen: "...entre más próximo a 0 es el valor del [índice de Gini], significa que existe una mayor concentración de los flujos en determinados puntos, y es el caso opuesto cuando el valor está próximo a 1." La interpretación del índice de Gini es justamente es al revés. El texto entre corchetes es nuestro.

<sup>12</sup> La explicación del índice Global de Gini se explica detalladamente en López-Vega y Velarde (2011) y en Plane y Mulligan (1997). En este segundo trabajo se presenta, incluso, un ejemplo de los cálculos "paso a paso".

del plano cartesiano a la esquina superior derecha. Esta es la llamada la *Línea de Igualdad* y corresponde a distribuciones perfectamente igualitarias

La CL es complemento *necesario* del índice de Gini, porque el índice registra una debilidad importante: puede generar el mismo valor para situaciones de desigualdad diferentes (Giorgi, 1993). Como el índice de Gini mide la *proporción* del área debajo de la Línea de Igualdad delimitada por la CL respecto al área total del triángulo que forman los ejes cartesianos y la Línea de Igualdad, puede producir un mismo valor para una curva obesa en la parte superior derecha de la curva y para una curva obesa en su parte inferior izquierda, que son escenarios muy distintos de desigualdad (Figura 1).

Figura 1. Curva Lorenz y su relación con el índice de Gini



Fuente: elaboración propia.

En la Figura 1 la CL “A” representa una distribución altamente *desigual* de migrantes entre las Entidades Federativas que registran *más* migrantes. Por su parte, la CL “B” también representa una distribución altamente *desigual* de migrantes entre las Entidades Federativas, pero en este caso respecto a las entidades con *menor cantidad* de migrantes. Sin embargo, como las áreas delimitadas por las CL son iguales (la CL “A” es el espejo de la CL “B”), el valor del índice de Gini es el mismo. Es decir: el Índice de Gini *no detecta las diferencias entre ambas situaciones*. Llama la atención

que la CL se utilice mucho menos que el índice de Gini en estudios migratorios, cuando son complementos mutuamente necesarios.

La *forma* de la CL es un buen indicador visual de desigualdad. Sin embargo, la inspección visual tiene diversas desventajas (Metzger, 2006) por lo que se requieren algunos indicadores cuantitativos que *sinteticen su perfil*. En este trabajo utilizamos tres:

La *Media*: donde la pendiente de la CL es igual a la de la Línea de Igualdad ( $p = 1.0$ ). Es el punto “A” en la Figura 2 que tiene como abscisa  $F(u)$ . Las entidades federativas que están a la derecha de ese punto son las *dominantes* en términos de migración: concentran más que proporcionalmente los movimientos migratorios (emigración o inmigración). El punto de quiebre  $F(u)$  es importante porque nos indica qué tan convexa es la CL. Mientras más convexa la curva, mayor desigualdad.

La *Mediana*. Indica cuántas entidades federativas concentran 50 por ciento de los movimientos migratorios. En la Figura 2 se observa que este punto (“B”) se localiza en la CL a partir de la abscisa 0.50 y del ángulo  $((Q0.5)/u)$ .

El *Índice de Schutz* (también conocido como de Pietra). Estima la máxima desviación entre la CL y la Línea de Igualdad. El valor del índice de Schutz se calcula de la siguiente forma (Figura 2):

$$F(\mu) - L(F(\mu)).$$

Estos tres indicadores sintetizan adecuadamente la *forma* de la CL, especialmente cuando las diferencias entre curvas son difíciles de apreciar visualmente (Lubrano, 2015) (Figura 2).

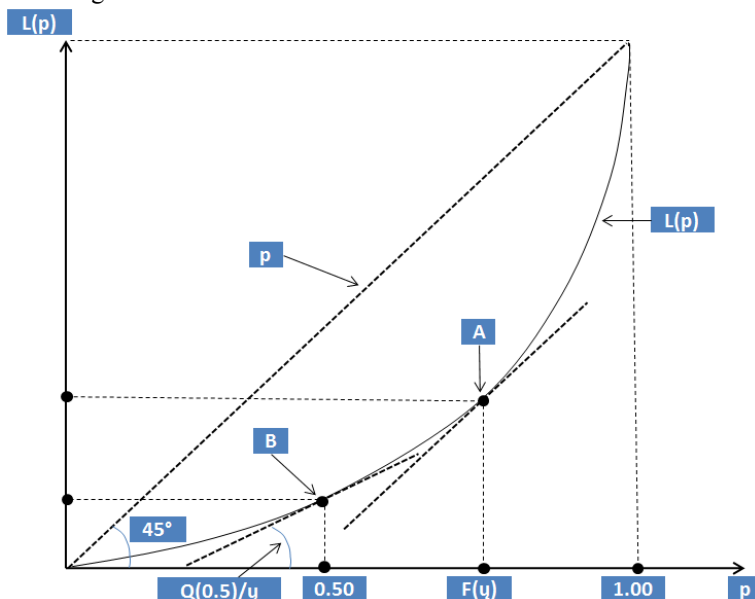
### Índice de Herfindahl-Hirschman

Debido a la debilidades ya mencionadas del índice de Gini, en este trabajo se utiliza también el índice de *Herfindahl-Hirschman* (IHH), que es una medida de concentración de segundo orden más poderoso que el índice de Gini (Masciandaro y Quintyn, 2009), aunque se recomienda usar ambos indicadores de manera complementaria (D’Amico y Di Biase, 2010).

El IHH ofrece una imagen de la *competencia en el mercado migratorio* a partir de la detección de unidades espaciales *dominantes* en mayor o menor medida: por ejemplo, entidades federativas que podrían dominar la atracción de inmigrantes o la expulsión de emigrantes. Al igual que el índice de Gini, el IHH utiliza las proporciones de los flujos que corresponden a cada entidad federativa, pero *elevadas al cuadrado*. Así se subraya el peso de las entidades más importantes en términos de migrantes y se reduce el

problema de generar valores iguales del IHH para situaciones diferentes, como sí ocurre con el índice de Gini (Giorgi, 1993).

Figura 2. Puntos clave de la forma de la Curva Lorenz



Fuente: Duclos y Araar, 2006

El IHH se calcula sumando los cuadrados de las participaciones del *mercado migratorio* de cada entidad federativa (v.g. Inmigración, Emigración y/o Total). Por ejemplo, si toda la inmigración llegara a una sola entidad federativa se diría que esa entidad concentra 100 por ciento del *mercado de inmigración* y no habría competencia. El IHH sería:  $100 \times 100 =$  diez mil que es el valor máximo del IHH. La situación extrema opuesta (también hipotética, evidentemente) sería que no hubiera inmigración y entonces todas las entidades federativas tendrían una participación del *mercado inmigratorio* igual a 0.0. En esta situación, el IHH sería 0.0 (v.g. la sumatoria de todos los cuadrados de cero). Por tanto, mientras más cercano a 10 mil sea el valor del IHH *menor es la competencia migratoria* y mientras más cercano a cero *mayor la competencia migratoria* (valores cercanos a cero indicarían, en términos tomados de la economía, una situación de *competencia migratoria casi perfecta*).

El IHH se calcula de la siguiente forma:

$$IHH = \sum_{i=1}^N s_i^2 \quad (6)$$

Donde

$$s_i = \frac{x_i}{\sum_{i=1}^N x_i}$$

es la suma de la fracción del total migratorio *por entidad federativa* al cuadrado de los datos que lo componen.

Adicionalmente, la literatura reporta *umbrales* del IHH para interpretar la competencia: valores inferiores a 1 500 indican *alta competencia*; valores entre 1 500 y 2 500 se interpretan como de *competencia moderada*; y de 1 800 y más se entienden como de *falta de competencia* (v.g. presencia de agentes altamente dominantes o monopolícos) (Sys, 2009, entre otros).<sup>13</sup>

### Coefficientes de Correlación de Pearson y Spearman

Las medidas de correlación de Pearson y Spearman son índices estadísticos que miden la relación lineal entre dos variables. Aquí se utilizan como medidas de *similitud del comportamiento estadístico*. La correlación de Pearson se estima a partir de la comparación de dos distribuciones de datos (ecuación 7). Por su parte, el coeficiente de correlación de Spearman se calcula a partir de la comparación de los rangos de datos apareados (ecuación 8). La correlación de Spearman puede ser calculada con la fórmula de Pearson si antes se agrupan los datos en rangos. Ambos coeficientes varían de -1 a +1 y se interpretan igual: si son iguales o cercanos a cero significa que las distribuciones de los flujos migratorios observados y calculados *no son similares*; si son cercanos o iguales a +1, significa que el comportamiento estadístico de las variables es *muy similar*; y si son iguales o cercanos a -1, significa que las observaciones se comportan en sentido perfectamente inverso: son completamente *disimilares*.

Los coeficientes de Pearson ( $r$ ) y Spearman ( $r_s$ ) se estiman de la siguiente manera:

$$r = \frac{S_{xy}}{S_x S_y} \quad (7)$$

<sup>13</sup> Lo que no sucede con el índice de Gini: ¿cuándo un valor es alto, medio o bajo? La interpretación del índice de Gini se hace en términos comparativos para cada caso de estudio.

Donde  $S_{xy}$  es la covarianza de flujo observado en relación con el flujo estimado,  $S_x$  y  $S_y$  es la desviación estándar de cada uno de los flujos.

$$r_s = 1 - \frac{6 \sum d^2}{k(k^2 - 1)} \quad (8)$$

Donde  $d$  es la diferencia entre los rangos (i.e. flujo observado menos flujo estimado),  $k$  es el número datos en el rango.

### Error proporcional promedio

El *error proporcional promedio* permite estimar la diferencia global promedio entre la matriz *observada* y las *calculadas*, comparando todos los flujos migratorios. Por tanto, permite dos cosas básicas en este texto: i. Identificar el orden ( $n$ ) del método de Medias Móviles que mejor se ajusta a los flujos observados. Es decir, el orden ( $n$ ) que minimiza el error proporcional promedio entre el sistema migratorio calculado con el modelo y el sistema migratorio *observado* (o *real*); y, ii. Estimar la desviación o falla proporcional promedio de las matrices calculadas mediante las CME y las CMD respecto de la matriz de flujos migratorios observados.

El error proporcional promedio ( $e$ ) se estima de la siguiente forma:

$$e = \frac{\sum |F_{ijo} - F_{ijc}|}{(2 \sum F_{ijc})} \quad (9)$$

donde:

$e$  = Error proporcional promedio de los flujos migratorios calculados (proporción de migrantes *erróneamente calculados* por el modelo):

$F_{ijo}$  = Estructura porcentual observada de los flujos migratorios.

$F_{ijc}$  = Estructura porcentual calculada de los flujos migratorios.

Vale subrayar que ( $e$ ) estima el error promedio de los cálculos respecto al parámetro de referencia, que en este trabajo fue la matriz de flujos observados de 2010, y los expresa en proporciones. Si se multiplican por 100 pueden ser interpretados como *porcentajes de error*.<sup>14</sup> El valor mínimo de ( $e$ ) es 0.0, esto implica que no existe diferencia entre los flujos calculados y la matriz de flujos *observados* para el mismo año: el modelo replica perfectamente lo observado. Sin embargo, esta situación es muy rara de encontrar, lo más usual es que exista un cierto error proporcional promedio

<sup>14</sup> Por ejemplo, ( $e$ ) = 0.20 significa que el modelo registra un error proporcional promedio de 20 por ciento respecto a los datos observados.



( $e$ ). Mientras más alto sea el valor de ( $e$ ), mayor es la diferencia entre lo calculado y lo observado: menos exacto es el modelo.

La valoración de la importancia de ( $e$ ) debe contrastarse con la naturaleza de cada situación que se está modelando y con el número de datos que proyecta el modelo. Si son pocos datos, un ( $e$ ) = 0.10 puede ser alto, si además se trata de un proceso delicado (relacionado con algún problema de salud, por ejemplo) podría ser inaceptable (Garrocho *et al.* 2002; Webber, 1984).<sup>15</sup> En el caso de este estudio sobre migración, un ( $e$ ) = 0.10 sería muy bueno, puesto que significaría que el modelo sólo falla sus proyecciones diez por ciento en promedio, aún y si está proyectando 992 datos (los flujos entre 32 estados, eliminando la diagonal de la matriz). También se debe recordar que ( $e$ ) es un indicador promedio *global* (como el índice de Gini global que se mencionó antes) y, por tanto, habrá casos (v.g. flujos entre entidades federativas) que registren errores o desviaciones mayores y menores al valor de ( $e$ ).

#### FUENTES DE INFORMACIÓN Y SOFTWARE UTILIZADO

En este trabajo se dispuso de una secuencia temporal de matrices de transición del sistema migratorio mexicano a escala de entidades federativas. Las matrices se construyeron con el apoyo del Consejo Nacional de Población (Conapo) utilizando los datos de *migración reciente* de los Censos de Población y Vivienda 1990, 2000 y 2010, y de los Conteos de Población 1995 y 2005.<sup>16</sup>

La migración reciente capta sólo el lugar de *residencia habitual* de la población cinco años antes de ser entrevistada para la construcción de las fuentes de información.<sup>17</sup> En el grupo de 0 a cuatro años se considera el lugar de nacimiento como el origen migratorio. Esta aproximación al registro de la migración reciente implica que no se captan los posibles cambios de lugar de residencia ocurridos en el quinquenio, lo que es una gran limitación de las fuentes de información migratoria de México, ya que no existen registros continuos de cambio de residencia (López-Vega y Velarde, 2011).

Adicionalmente, las fuentes de información de México consideran como cambios de residencia sólo aquellos que implican cruzar límites estatales o municipales. Por tanto, en la migración interestatal no se registra

<sup>15</sup> La valoración de qué tan aceptable es el valor de ( $e$ ) no responde a reglas preestablecidas, sino a las características del estudio que se esté realizando. Algo similar ocurre con otros indicadores estadísticos como el índice de Gini o el coeficiente de correlación de Pearson ( $R$ ): ¿Cuándo un coeficiente de correlación ( $R$ ) es alto, mediano o bajo? (Field, 2005).

<sup>16</sup> Como lo hacen López-Vega y Velarde en su espléndido trabajo de 2011.

<sup>17</sup> Véase una amplia discusión sobre lo difuso del término *residencia habitual* en Garrocho, 2011.

un cambio de residencia, aunque involucre un cambio profundo para la población (por distancia, por cambio cultural...), si ocurrió dentro de la misma entidad federativa (por ejemplo una familia que migra de la ciudad de San Luis Potosí a la Huasteca potosina). Sin embargo, un cambio de residencia que implica moverse una cuadra (que podría ser, teóricamente, de menos de cincuenta metros), pero que implicó un cambio de entidad federativa sí se registra como migración interestatal (por ejemplo, una familia que cambia su residencia en la misma colonia, como la Lomas de Sotelo, que tiene una parte en el Estado de México y otra en el Distrito Federal). Evidentemente, se requieren datos más detallados y precisos sobre las migraciones en nuestro país (Garrocho, 2011).

Estas limitaciones son *inherentes* a las fuentes de información de México. Sin embargo, aun así es posible sintetizar razonablemente los datos de los flujos migratorios interestatales en matrices origen-destino (MOD) de 32 renglones por 32 columnas. La MOD consigna el intercambio de población entre diferentes unidades espaciales (e.g. estados, municipios). Se lee de derecha a izquierda y de arriba abajo. En cada celda de la matriz se registran los flujos de *salida* (si se lee por renglón) o *llegada* (si se lee por columna) de migrantes interestatales. La suma de cada renglón es el total de población que salió *desde* cada entidad. La suma de cada columna es el total de población que llegó *a* cada entidad. La diagonal de la matriz representa la población que no migró. La celda de la matriz ubicada en el cruce de la columna y el renglón de las sumas (v.g. la que registra simultáneamente los totales por renglón y por columna) es el total de población que migró en el periodo considerado. La MOD de migración interestatal para México se compone de 1 024 casillas ( $32 \times 32 = 1\,024$ ). Si se descuenta la diagonal (32 casillas: la población que no migró) resultan 992 flujos migratorios interestatales. Con esta información se estimaron en este texto las Cadenas de Markov Estáticas y Dinámicas. Los cálculos se realizaron con MatLab, principalmente.

### **CADENAS DE MARKOV ESTÁTICAS *VERSUS* CADENAS DE MARKOV DINÁMICAS CON MEDIAS MÓVILES**

#### **Estrategia de comparación de las Cadenas de Markov Estáticas (CME) *versus* las Cadenas de Markov Dinámicas con Medias Móviles (CMD)**

La estrategia que se siguió para comparar la precisión de las estimaciones de las CME *versus* las CMD fue la siguiente:

- a) La matriz de flujos *observados* de 2010 (la más reciente disponible) fue el parámetro de referencia que se utilizó para valorar la *precisión / imprecisión* de las estimaciones de las CME y las CMD. Especialmente el renglón y la columna de totales de la matriz.
- b) Las *Métricas de Bondad de Ajuste* entre la matriz de flujos *observados* (parámetro de referencia) y los valores *calculados* por las Cadenas de Markov Estáticas y Dinámicas permitieron estimar las cuatro dimensiones de los resultados: Desigualdad, Competencia Migratoria, similitud Estadística y Error Global. En este trabajo se considera que las tres primeras dimensiones tienen la *misma importancia* (una estrategia similar a la de López-Vega y Velarde, 2011) y que el error proporcional promedio es uno de los indicadores de bondad de ajuste *más poderosos* por su carácter global (como lo consideran desde Webber, 1984, hasta: Boutros *et al.*, 2015; Chormanski *et al.*, 2008 o Singh, 2001, entre muchos otros)
- c) Las CME para 2010 se estimaron considerando las matrices de migración de 2000 y de 2005.
- d) Las CMD se computaron considerando las siguientes ventanas temporales: *i.* Las matrices de 1990 y 2000 ( $n = 2$ ); *ii.* Las matrices de 2000 y 2005 ( $n = 2$ ); *iii.* Las matrices de 1995, 2000, 2005 ( $n = 3$ ); *iv.* Las matrices de 1990, 1995, 2000 ( $n = 3$ ); y, *v.* Las matrices de 1990, 1995, 2000, 2005 ( $n = 4$ ).

### **Bondad de Ajuste en las cuatro dimensiones consideradas: desigualdad, competencia migratoria, similitud estadística y error proporcional promedio**

#### *Dimensión 1: Desigualdad*

Emigración. En el año 2000 el Índice de Gini (IG) de *emigración total observada* por entidad federativa registró un valor de apenas 0.10, lo que significa baja desigualdad interestatal en términos de emisión de población. Esta situación permanece prácticamente igual en 2010: el IG es 0.11 (Cuadro 1). Las CME replican bien la desigualdad de la emigración observada en 2000 (con un IG de 0.08) y también muestran un escenario sin cambios importantes de desigualdad para 2010 (como ocurre con la situación observada). En cambio, las CMD generan escenarios más alejados de los observados: el IG de 2000 es 0.02 y el 2010 es 0.01, mucho menos desiguales que los observados y que los estimados por las CCE. Aunque las CME y las CMD generan escenarios de desigualdad emigratoria baja

y estable en el periodo (que es el mismo comportamiento que los flujos observados), las CME se aproximaron más a la situación observada. Punto a favor de las CME.

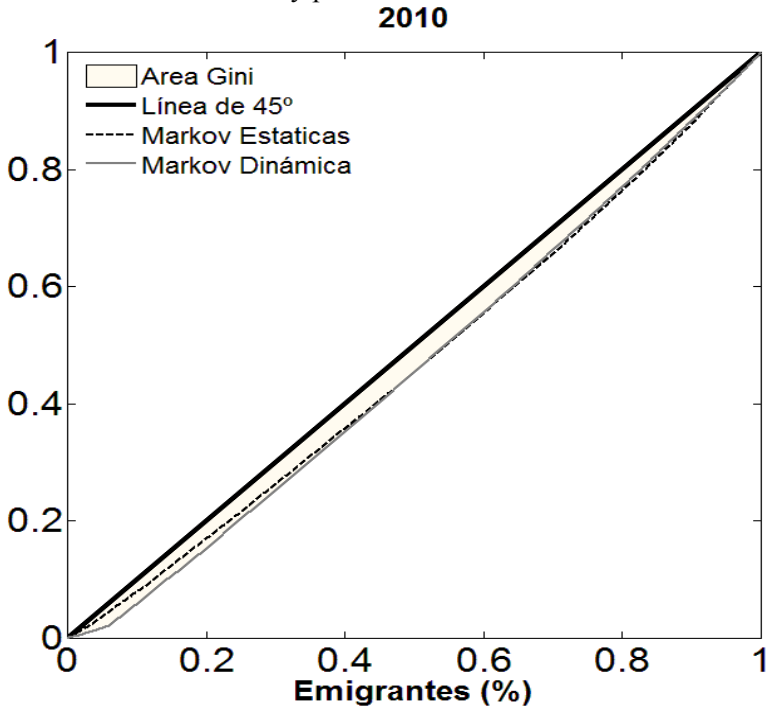
Por su parte, los indicadores de la *forma* de las Curvas de Lorenz (CL) muestran que las curvas generadas por ambos modelos markovianos son muy parecidas a la que resulta de los valores observados, lo que es consistente con el análisis visual que no permite detectar a simple vista diferencias entre ellas (Figura 3; ver las columnas *Diferencias* en el Cuadro 1). La distancia euclidiana entre el punto de *Emigración Media* de la CL observada y el que producen las CME para 2000 es apenas 0.02 y 0.04 para las CMD; las diferencias en la *Mediana* de las CL de las CME y las CMD respecto a la *Mediana* de la CL observada son 0.00 y 0.01, respectivamente; y las diferencias entre los Índices de Schutz también son marginales: 0.09 y 0.05.

También en 2010, las diferencias entre la forma de la CL observada y las de las CL calculadas fueron bajas: la *Media* de la CL derivada de los cálculos de la CME (curva que abreviaremos como CLE) se desvió tres centésimas de la Media de la CL observada y la Media de la CL perfilada con cálculos de la CMD (que abreviaremos como CLD) falló por sólo cuatro centésimas. La *Mediana* de la CLE se desvió nueve centésimas de la localización de la Mediana de la CL observada y la CLD erró por sólo dos centésimas (Cuadro 2).

También se observa una gran similitud de la forma de las CL cuando se analizan los puntos de máxima desviación entre las CL y la Línea de Igualdad, que se estima mediante el Índice de Schutz. La diferencia entre la CL observada y la CLE es de sólo dos centésimas y de siete centésimas respecto a la CLD (Cuadro 2). Llama la atención la gran capacidad de los modelos Markovianos para replicar la forma de la CL perfilada con los datos observados. Para efectos prácticos, las CL calculadas y la observada son iguales. Por tanto, en términos de la forma de las CL declaramos *empate*.

Inmigración. En 2000 la desigualdad de inmigración total observada interestatal es baja: registra un IG de 0.10, que desciende ligeramente para 2010, cuando el IG se reduce a 0.06, lo que indica una reducción de la desigualdad inmigratoria entre las entidades federativas (Cuadro 1). Las CME logran replicar bien la situación observada en 2000 ya que registran un IG de los totales estatales de inmigración igual a 0.06. Para 2010 incrementan su precisión: mantienen su IG en 0.06 con lo que la desigualdad que replican en el sistema migratorio es igual a la observada.

Figura 3. Imposibilidad de comparar visualmente las Curvas de Lorenz cuando son muy parecidas



Fuente: elaboración propia.

En cambio, las CMD fallan en 2000 con un IG de 0.01, menor al observado y al estimado por las CME, y aunque en 2010 se acercan notablemente a la situación observada con un IG de 0.04, no logran la perfecta exactitud de las CME. Otro punto para las CME.

Como ocurrió con la emigración, la forma de la CL observada de *inmigración* es bien replicada con los datos calculados por los modelos de Markov. En 2000, el punto de inmigración *Media* (que permite identificar las entidades que concentran más que proporcionalmente los movimientos inmigratorios) de la CLE se desvía sólo ocho centésimas de la Media de la CL observada (recordar que la escala de la CL va de 0.0 a 1.0) y la Media de la CLD sólo falla dos centésimas (casi coincide perfectamente con la Media de la CL observada). Las diferencias en la localización de la *Mediana* (que indica cuántas entidades federativas concentran 50 por ciento de los movimientos migratorios) de la CLE falló por sólo una centésima respecto a la Mediana de la CL observada y la CLD se desvió sólo tres centésimas.

Cuadro 1. Indicadores de Bondad de Ajuste entre los flujos migratorios interestatales observados y calculados por emigración, inmigración y global, 2000-2010

Indicadores de Bondad de Ajuste	2000			2010		
	Observados	Estimados		Observados	Estimados	
		CM Estáticas	CM Dinámicas		CM Estáticas	CM Dinámicas
<i>Índice de Gini</i>						
Emigración	0.10	0.08	0.02	0.11	0.07	0.01
Inmigración	0.10	0.06	0.01	0.06	0.06	0.04
Global	0.24	0.22	0.21	0.24	0.24	0.23
<i>Índice de Herfindhal-Hirschman</i>						
Emigración	492.8	714.7	774.8	455.2	413.0	442.7
Inmigración	455.2	621.7	754.7	485.2	779.8	639.9
Global	461.2	741.1	787.2	714.7	731.2	954.7
<i>Índice de Correlación de Pearson</i>						
Emigración	N.A.	0.35	0.37	N.A.	0.36	0.39
Inmigración	N.A.	0.45	0.33	N.A.	0.35	0.35
Global	N.A.	0.46	0.41	N.A.	0.37	0.45
<i>Índice de Correlación de Spearman</i>						
Emigración	N.A.	0.25	0.40	N.A.	0.30	0.41
Inmigración	N.A.	0.20	0.45	N.A.	0.28	0.38
Error proporcional promedio	N.A.	0.06	0.04	N.A.	0.04	0.02

Fuente: elaboración propia.

Cuadro 2. Indicadores de la forma de las Curvas de Lorenz, 2000-2010

Indicadores	2000									
	Cadenas de Markov Estáticas					Emigración				
	Cordenadas		Pendiente	L(p)	Q(x)/ $\mu$	Cordenadas		Pendiente	L(p)	Q(x)/ $\mu$
Emigración media	0.43	0.39	1.00	45.00	0.38	0.37	1.00	45.10		
Punto de concentración de 50%	0.50	0.46	2.61	69.0	0.50	0.45	2.80	70.30		
Punto de máxima desigualdad	0.46	0.45	3.06	71.90	0.50	0.45	2.28	70.30		
	Inmigración									
	Cordenadas		Pendiente	Ángulo(°)	Cordenadas		Pendiente	Ángulo(°)		
Emigración media	0.13	0.11	1.01	45.20	0.09	0.07	1.01	45.30		
Punto de concentración de 50%	0.50	0.47	0.69	34.80	0.50	0.49	1.32	52.90		
Punto de máxima desigualdad	0.36	0.34	1.61	58.10	0.34	0.32	1.42	55.00		
	2010									
	Cordenadas		Pendiente	Ángulo(°)	Cordenadas		Pendiente	Ángulo(°)		
Emigración media	0.13	0.11	0.99	44.90	0.07	0.08	1.00	45.10		
Punto de concentración de 50%	0.50	0.40	0.58	30.20	0.50	0.47	1.25	60.30		
Punto de máxima desigualdad	0.41	0.38	2.29	66.40	0.41	0.47	2.14	64.90		
	Inmigración									
	Cordenadas		Pendiente	Ángulo(°)	Cordenadas		Pendiente	Ángulo(°)		
Emigración media	0.36	0.35	1.00	45.00	0.39	0.38	1.00	45.00		
Punto de concentración de 50%	0.51	0.49	0.62	32.00	0.50	0.43	1.11	48.10		
Punto de máxima desigualdad	0.28	0.27	3.03	71.70	0.24	0.20	5.09	78.90		

Fuente: elaboración propia.

Cuadro 2. Indicadores de la forma de las Curvas de Lorenz, 2000-2010 (continuación)

Indicadores	2000									
	Observados					Emigración				
	Cordenadas		Pendiente	L(p)	Q(x)/ $\mu$	Cordenadas		Ángulo(°)	Diferencia	
Emigración media	x	y		L(p)	Q(x)/ $\mu$			Distancia Estática	Distancia Dinámica	
Punto de concentración de 50%	0.41	0.39	1.00	1.00	45.0			0.02	0.04	
Punto de máxima desigualdad	0.50	0.46	0.90	2.61	42.4			0.00	0.01	
	0.55	0.47	2.61		69.1			0.09	0.05	
	Inmigración									
	Cordenadas		Pendiente	L(p)	Q(x)/ $\mu$	Cordenadas		Ángulo(°)	Diferencia	
Emigración media	x	y		L(p)	Q(x)/ $\mu$			Distancia Estática	Distancia Dinámica	
Punto de concentración de 50%	0.07	0.06	1.00	1.00	45.00			0.08	0.02	
Punto de máxima desigualdad	0.50	0.46	2.41	2.74	67.50			0.01	0.03	
	0.36	0.34	2.74		69.90			0.00	0.03	
	2010									
	Observados					Emigración				
	Cordenadas		Pendiente	L(p)	Q(x)/ $\mu$	Cordenadas		Ángulo(°)	Diferencia	
Emigración media	x	y		L(p)	Q(x)/ $\mu$			Distancia Estática	Distancia Dinámica	
Punto de concentración de 50%	0.11	0.09	1.00	1.00	45.00			0.03	0.04	
Punto de máxima desigualdad	0.50	0.49	0.66	2.21	33.63			0.09	0.02	
	0.41	0.40	2.21		65.66			0.02	0.07	
	Inmigración									
	Cordenadas		Pendiente	L(p)	Q(x)/ $\mu$	Cordenadas		Ángulo(°)	Diferencia	
Emigración media	x	y		L(p)	Q(x)/ $\mu$			Distancia Estática	Distancia Dinámica	
Punto de concentración de 50%	0.37	0.35	1.00	1.00	45.00			0.01	0.04	
Punto de máxima desigualdad	0.50	0.49	2.95	5.09	71.31			0.01	0.06	
	0.24	0.20	5.09		78.90			0.08	0.00	

Fuente: elaboración propia.



Por su parte, el punto de máxima desviación entre la CL y la Línea de Igualdad (calculado mediante el Índice de Shutz) también registra coincidencias altas entre la CL observada y las calculadas: se ajusta perfectamente en el caso de la CLE (la desviación es 0.00) y la CLD falla por sólo tres centésimas respecto a la CL observada. En 2010, los ajustes son también elevados. La *Media* de la CLE falla por sólo una centésima respecto a la CL observada y la CLD se desvía apenas cuatro centésimas. La *Mediana* y el punto de máxima desviación entre la CL y la Línea de Igualdad de las CL muestran una situación muy similar: desviaciones de una y seis centésimas para el caso de la Media y de ocho centésimas a un ajuste perfecto (el de la CLD) para el caso del Índice de Shutz. Ambos modelos markovianos replican igual de bien la forma de la CL observada: empate.

Desigualdad migratoria global. La desigualdad global observada en el sistema migratorio en 2000 (estimada con el método de Plane y Mulligan, 1997) puede calificarse de elevada ya que el IG es 0.24 y permanece igual en 2010 (Cuadro 1). Las CME reflejan bien esta gran desigualdad ya que su IG para 2000 es 0.22 y justo de 0.24 en 2010, lo que replica casi exactamente la desigualdad promedio que se registra en la situación observada. Las CMD muestran un desempeño muy parecido al de las CME: IG = 0.21 en 2000 y 0.23 en 2010, por lo que podemos declarar empate en este rubro. La desigualdad migratoria con el método de Plane y Mulligan (1997) no permite derivar Curvas de Lorenz.

### *Dimensión 2: Competencia Migratoria*

Emigración. En general, ambos modelos markovianos replican relativamente bien la competencia emigratoria interestatal observada para 2000 y 2010 (recordemos que la escala del IHH es de 0 a 10,000) (Cuadro 1). El IHH observado en 2000 es 492.8, el de las CME es 714.7 y el de las CMD de 774.8. Para 2010 estos valores permanecen estables (455.2, 413.0 y 442.7, respectivamente). Los valores cercanos a cero del IHH indican que la competencia emigratoria observada y calculada es muy alta, lo que implica que no existen entidades federativas dominantes como expulsoras de población. Esto es coherente con la evidencia proporcionada por el IG. En el tema de la emigración medida con el IHH podemos declarar empate entre las CME y las CMD: las diferencias de sus IHH respecto al de emigración observada son marginales, ya que la escala de medición va de 0.0 a 10 mil.

Inmigración. La situación en este rubro es muy similar a la de emigración: los dos modelos de Markov reproducen con exactitud la competencia

inmigratoria observada a escala interestatal tanto en 2000 como en 2010 (Cuadro 1). Los valores de los IHH en 2000 van de 455.2 a 754.7 y en 2010 el rango es de 485.2 a 779.8. Es decir, la competencia inmigratoria se mantuvo estable. Los valores cercanos a cero indican que no existen entidades dominantes en materia de inmigración, sino que la competencia inmigratoria es muy alta. El comportamiento estable de la competencia inmigratoria en los años estudiados también es consistente con lo que muestra el IG. Dado que las diferencias en los IHH de las CME y las CMD respecto al de inmigración observada son marginales, podemos declarar empate entre los modelos markovianos.

Competencia Migratoria Global. Como es de esperarse (si se revisaron los resultados anteriores), la competencia migratoria global es muy alta tanto en la situación observada, como en los resultados de los modelos de Markov: los IHH varían en rangos estrechos que van de 461.2 a 787.2 en 2000 y de 714.7 a 954.7 en 2010 (Cuadro 1). Los IHH de las CME se acercan un poco más a valor del IHH de la situación observada que los de las CMD, pero estas diferencias son tan pequeñas (en una escala de 0.0 a 10 mil) que podemos afirmar que para efectos prácticos la exactitud de ambos modelos es igual y se puede declarar empate. En este rubro los resultados no son directamente comparables con los del IG, que mostraron desigualdad global creciente, porque la *desigualdad global* se estimó con el Método de Plane y Mulligan (1997).

### *Dimensión 3: similitud del comportamiento estadístico*

Emigración. Los coeficientes de correlación de Pearson entre los valores calculados y los observados tienen en todos los casos signo positivo y son de magnitud muy similar, tanto para 2000 como para 2010. En el primer año registran 0.35 (CME) y 0.37 (CMD) y en el segundo 0.36 (CME) y 0.39 (Cuadro 1). Si consideramos los umbrales de Field (2005), puede decirse que los dos modelos markovianos reproducen con precisión *intermedia* el comportamiento estadístico de la emigración interestatal observada.<sup>18</sup> En ambos años las CMD registran índices de correlación ligeramente más elevados que los de las CME, especialmente en 2010. Dado el tamaño de la muestra (32 totales estatales) las diferencias no son marginales. Esto se confirma si revisamos los coeficientes de correlación de rangos de Spearman: las CMD generan coeficientes más elevados que las CME. En 2000 y 2010

<sup>18</sup> Field (2005: 112) se hace la siguiente pregunta: ¿Qué tan alta es la correlación entre dos variables? y luego de una amplia revisión de la literatura propone tres *umbrales guía* (todos con signos positivo y negativo): un coeficiente alrededor de 0.1 es *Bajo*; alrededor de 0.30 *Medio*; y uno igual o mayor de 0.50 sería *Alto*.

los coeficientes de Spearman entre los totales observados y los de las CMD fueron 0.40 y 0.41, mientras que los de las CME llegaron a 0.25 en 2000 y 0.30 en 2010. Punto para las CMD.

Inmigración. En 2000 las CME registran un coeficiente de correlación de Pearson con el patrón observado mucho más alto que las CMD: 0.45 versus 0.33, pero en 2010 ambos modelos producen el mismo coeficiente: 0.35 (Cuadro 1). Sin embargo, es interesante que el coeficiente de Spearman es mucho más alto para las CMD: en 2000 es 0.45 y el de las CME es 0.20, y en 2010 los valores son 0.38 y 0.28 respectivamente. En otras palabras, la tendencia de los valores individuales es mejor reflejada por las CME, pero la importancia (v.g. *jerarquía*) de las entidades federativas como atractoras de población la reproducen mejor las CMD. En esta situación los propósitos de cada estudio definirán cuál modelo markoviano es más adecuado, por lo que en este rubro declaramos empate.

Similitud estadística global. Lo primero que se debe resaltar es que los coeficientes de correlación de Pearson muestran valores intermedios y altos, lo que indica que ambos modelos replican bien los valores observados (Cuadro 1). En 2000 los coeficientes de los dos modelos son parecidos (CME = 0.46 y CMD = 0.41), pero en 2010 las CMD superan claramente a las CME: los valores son 0.45 y 0.37, respectivamente. En términos de similitud estadística parece que mientras más información se tenga de la tendencia (v.g. de la trayectoria) mejor es el rendimiento de las CMD. Punto para las CMD. Para el caso de la similitud estadística global no es posible calcular el coeficiente de correlación de Spearman.

#### *Dimensión 4: Error Proporcional Promedio*

Los dos modelos markovianos muestran una notable capacidad de modelar los flujos inmigratorios interestatales: considerando las 992 casillas de la matriz de flujos migratorios registran un error proporcional promedio ( $e$ ) máximo de seis por ciento y mínimo de dos por ciento (Cuadro 1). En otras palabras: los modelos aciertan sus réplicas del patrón observados 94 por ciento en promedio, como mínimo, pero pueden llegar a registrar una precisión promedio de 98 por ciento. Estos resultados sorprendentes los verificamos varias veces: son correctos. Para ambos años de estudio las CMD registran ( $e$ ) más bajos y aunque la diferencia no es notable, otorgamos el punto a las CMD por su mayor exactitud en este *indicador calve* de bondad de ajuste global.

## CONCLUSIONES Y APORTACIONES

En este trabajo se realizó una comparación sistemáticamente *multidimensional* de los modelos de Markov estáticos y dinámicos con medias móviles, a fin de evaluar su precisión / imprecisión para replicar los flujos migratorios interestatales observados de México en 2000 y 2010. Se evaluaron cuatro dimensiones de los resultados: i) *Desigualdad* interestatal en la distribución de los flujos migratorios; ii) *Competencia* migratoria entre las entidades federativas; iii) *Similitud estadística* del comportamiento de los flujos entre estados; y, iv. *Error proporcional promedio* de los flujos estimados respecto a los observados.

Esta estrategia del análisis comparativo registra, al menos, tres particularidades metodológicas que vale subrayar. *Primero*, analiza dimensiones diferentes de los resultados para lograr dos objetivos: evaluar de manera más completa la precisión / imprecisión de los modelos markovianos y evitar “la paradoja de los relojes desincronizados” (¿cuál marca la hora correcta?). Es decir, en este trabajo no se mide la misma dimensión de los flujos migratorios de diferente manera, sino que se analizan aspectos diferentes de la misma red migratoria. *Segundo*, la desigualdad migratoria se analizó no sólo con índices de Gini, como normalmente se hace en los estudios sobre la distribución espacial de los flujos migratorios (v.g. *spatial focus*), sino que se apoyó en un estudio profundo de *la forma* de las Curvas de Lorenz usando indicadores clave. Dado que el índice de Gini puede generar el mismo valor para escenarios de desigualdad muy distintos, el uso de las Curvas de Lorenz es fundamental, aunque (sorprendentemente) rara vez se hace en los análisis de flujos migratorios. *Tercero*, el análisis de la competencia migratoria se apoyó en índices de Herfindahl-Hirschman, que son ampliamente utilizados en economía, pero que, inexplicablemente, poco se aplican en estudios sobre el grado de concentración de la migración entre unidades espaciales de origen y destino.

En términos analíticos, llama la atención la gran capacidad de los modelos Markovianos para replicar no sólo la red de flujos migratorios en sus distintas dimensiones, sino incluso la forma de las Curvas de Lorenz (en el caso de la desigualdad migratoria interestatal). Aún más, los resultados de los modelos estático y dinámico siempre apuntan en la misma dirección, evitan “la paradoja de los relojes desincronizados” y son consistentes en todas las dimensiones de sus resultados. Lo único que varía es su grado de precisión (siempre elevada) en alguna de las dimensiones. Un dato contundente: el indicador global de bondad de ajuste entre lo calculado y lo ob-

servado muestra que los modelos aciertan sus réplicas de la red migratoria observada 94 por ciento en promedio como mínimo (cuando se utiliza el modelo estático) y pueden llegar a registrar una precisión promedio de 98 por ciento si se aplica el modelo dinámico.

Al final, el modelo de Markov dinámico con medias móviles resultó ligeramente más preciso que el modelo estático para replicar los flujos migratorios interestatales observados de México (Cuadro 3). Tres argumentos justifican esta conclusión: i) mayor precisión global (menor error proporcional promedio); ii) exactitud notablemente superior para replicar el comportamiento estadístico de los flujos entre estados; y, iii) su capacidad para integrar mayor información al modelaje de los flujos migratorios (incorpora no sólo el presente, sino el pasado: las tendencias o trayectorias), lo que implica que mientras más amplia sea la temporalidad del análisis más información se incorpora sobre la tendencia migratoria y mayor la superioridad del modelo dinámico sobre el estático. Esta no es una especulación teórica: a lo largo del trabajo se observó cómo el modelo dinámico mejoró notablemente su exactitud en 2010 respecto a 2000.<sup>19</sup>

Sin embargo, los modelos markovianos estáticos y dinámicos siguen siendo vulnerables a su crítica principal: son buenas herramientas para modelar flujos migratorios, pero *no explican el fenómeno* que modelan. En este trabajo replican de manera notable los flujos migratorios interestatales de México, pero no explican *por qué* se generan esos flujos. Explicarlos sería materia de otros enfoques demográficos complementarios. Pedirles explicaciones a los modelos markovianos sería pedirles demasiado.<sup>20</sup>

<sup>19</sup> La superioridad de los modelos dinámicos está en línea con los hallazgos de Guijarro y Hierro (2005) para el sistema migratorio de España.

<sup>20</sup> El tema de la explicación en las ciencias sociales implica un debate epistemológico de amplias proporciones. Incluso existen diversas *teorías de la explicación*. En ciencias sociales se acepta que no siempre es posible generar explicaciones causales de la manera como se hace en los campos de las ciencias de la materia (como la física o la química) o en las ciencias de la vida (como la biología), lo que constituye un déficit epistemológico inherente de las ciencias sociales (Rappaport, 1995).

Cuadro 3. Síntesis: Cadenas de Markov Estáticas (CME) *versus* Cadenas de Markov Dinámicas (CMD)

Dimensiones a evaluar	Cadenas de Markov	
	Estáticas	Dinámicas
<i>Desigualdad</i>		
Índice de Gini		
Emigración	+	
Inmigración	+	
Global	=	=
<i>Forma de la Curva de Lorenz</i>		
Emigración	=	=
Inmigración	=	=
<i>Competencia Migratoria</i>		
Emigración	=	=
Inmigración	=	=
Global	=	=
<i>Similitud Estadística</i>		
Emigración		+
Inmigración	=	=
Global		+
Error Proporcional Promedio		+

Fuente: elaboración propia.

## BIBLIOGRAFÍA

ALESSIO, E., Carbone; A. G. CASTELLI y V. FRAPPIETRO, 2002, "Second-order moving average and scaling of stochastic time series", in *The European Physical Journal B-Condensed Matter and Complex Systems*, vol. 27, núm. 2, pp. 197-200.

ARAAR, Abdelkrim y Jean-Yves DUCLOS, 2006, *Poverty and equity: measurement, policy and estimation with DAD*, Springer-Verlag, NY.

- BELL, M., Charles-Edwards; E., KUPISZEWSKA, D., KUPISZEWSKI, M., STILLWELL, J. y Y. ZHU, 2015, "Internal migration data around the world: Assessing contemporary practice", in *Population, Space and Place*, vol. 21, núm. 1, pp. 1-17.
- BOUTROS, N.; M.R. SHORTIS y E.S. HARVEY, 2015, "A comparison of calibration methods and system configurations of underwater stereo video systems for applications in marine ecology", in *Limnology and Oceanography: Methods*, vol. 13, núm. 5, pp. 224-236.
- BUSTOS, A., 2013, "Estimación markoviana del tamaño y de las estructuras de una población", en *Realidad, Datos y Espacio Revista Internacional de Estadística y Geografía*, vol. 4, núm. 1, pp. 122-159.
- CASTLES, S.; H. DE HAAS, y M.J. MILLER, 2013, *The age of migration: international population movements in the modern world*, Guildford Press. New York-London.
- CHAKRABORTI, A. y G. GERMANO, 2010, "Agent-based models of economic interactions", in G. Naldi, L. Pareschi y G. Toscani, *Mathematical modeling of collective behavior in socio-economic and life sciences*, Springer Science & Business Media, pp. 3-29. NY.
- CHORMANSKI, Jaroslaw; Tim VAN DE VOORDE, Tim DE ROECK, Okke BATELAAN y Frank CANTERS, 2008, "Improving distributed runoff prediction in urbanized catchments with remote sensing based estimates of impervious surface cover", in *Sensors*, vol. 8, núm. 2, pp. 910-932.
- CHING, W.K., X. HUANG, M. NG, y T.K. SIU, 2013, *Markov Chains Models, Algorithms and Applications*, Springer. NY.
- CONT, R. y L. WAGALATH, 2013, "Running for the exit: distressed selling and endogenous correlation in financial markets", in *Mathematical Finance*, vol. 23, núm. 4, pp. 718-741.
- D'AMICO, Guglielmo y Giuseppe DI BIASE, 2010, "Generalized concentration/inequality indices of economic systems evolving in time", in *Wseas Transactions on Mathematics*, vol. 9, núm. 2, pp. 140-149.
- DE CASTRO, M.C., 2007, "S"patial demography: An opportunity to improve policy making at diverse decision levels", in *Population research and policy review*, vol. 26, núm. 5-6, pp. 477-509.
- FIELD, A., 2005, *Discovering Statistics Using SPSS*, Sage Publications. Londres.
- GARCÍA-ABAD, R., 2003, "Un estado de la cuestión de las teorías de las migraciones", en *Historia Contemporánea*, núm. 26, pp. 329-351.
- GARROCHO, C., 2011, *Población flotante, población en movimiento: conceptos clave y métodos de análisis exitosos*, United Nations Population Fund-El Colegio Mexiquense-Conapo. México.
- GARROCHO, C., T. CHÁVEZ, y J.A. ÁLVAREZ-LOBATO, 2002, *La dimensión espacial de la competencia comercial*, El Colegio Mexiquense. México.

- GEYS, B., 2006, "Looking across borders: A test of spatial policy interdependence using local government efficiency ratings", in *Journal of Urban Economics*, vol. 60, núm. 3, pp. 443-462.
- GIMENEZ, O., LEBRETON, J. D., GAILLARD, J. M., CHOQUET, R., y R. PRADEL, 2012, "Estimating demographic parameters using hidden process dynamic models", in *Theoretical population biology*, vol. 82, núm. 4, pp. 307-316.
- GIORGI, G.M., 1993, "A fresh look at the topical interest of the Gini concentration ratio", in *Metron*, vol. 51, núm. 1-2, pp. 83-98.
- GÓMEZ-GARCÍA, J., M.A. PALACIOS y J. GARCÍA-PÉREZ, 1997, "Movimientos migratorios intermunicipales en la Comunidad Autónoma de Murcia: un enfoque markoviano", en *Cuadernos de Economía Murciana*, vol. 21, núm. 1, pp. 61-73.
- GOODKIND, D. y L.A. WEST, 2002, "China's floating population: definitions, data and recent findings", in *Urban Studies*, vol. 39, núm. 12, pp. 2237-2250.
- GUIJARRO, M. y M. HIERRO, 2007, "Un análisis global y desagregado de la dispersión espacial de las migraciones interiores en España (1986-2003)", en *Cuadernos Aragoneses de Economía*, vol. 17, núm. 1, pp. 163-180.
- GUIJARRO, M. y M. HIERRO, 2005, "Un análisis de la dinámica de los movimientos migratorios interregionales en España (1986-2001): una explotación del método MCC", en *Investigaciones Regionales*, núm 6, pp. 125-140.
- HARVIR, S.B., S.F. TAYLOR y Y. ST. JAMES, 2005, "Migrating to new service providers: toward a unifying framework of consumers' switching behaviors", in *Journal of the Academy of Marketing Science*, vol. 33, núm. 1, pp. 196-115.
- HIERRO, M. y M. GUIJARRO, 2006, "Un estudio mediante cadenas de Markov de la dinámica de los movimientos migratorios interterritoriales en España (1990-2003) desde un planteamiento de estimación dinámico", en *Revista Asturiana de Economía*, núm. 35, pp. 145-161.
- HIERRO, M. y M. GUIJARRO, 2007, "Una revisión de la aplicación de las cadenas de Markov discretas al estudio de la movilidad geográfica", en *Estadística española*, vol. 49, núm. 166, pp. 473-499.
- KAYACAN, E., B. ULUTAS, y O. KAYNAK, 2010, "Grey system theory-based models in time series prediction", in *Expert Systems with Applications*, vol. 37, núm. 2, pp. 1784-1789.
- LE GALLO, Julie y Sandy DALL'ERBA, 2006, "Evaluating the temporal and spatial heterogeneity of the European convergence process, 1980-1999", in *Journal of Regional Science*, vol. 46, núm. 2, pp. 269-288.
- LIN, S.N., C.Y. CHOU, S.L. WANG y L. HUI-RONG, 2012, "Economic design of autoregressive moving average control chart using genetic algorithms", in *Expert systems with applications*, vol. 39, núm. 2, pp. 1793-1798.
- LÓPEZ-VEGA, R. y S. VELARDE, 2011, "Aplicación de medidas de concentración para el análisis demográfico de la migración interna de México", en *La*



situación demográfica de México 2011, Consejo Nacional de Población, CDMX, pp. 123-139.

LUBRANO, M., 2015, *The econometrics of inequality and poverty: Lorenz curves, the Gini Coefficient and parametric distributions*, Le Groupement de Recherche en Économie Quantitative d'Aix-Marseille (GREQAM), Marsella, Francia.

MASCIANDARO, D. y M. QUINTYN, 2009, "Measuring financial regulation architectures and the role of the central banks: the financial supervision Herfindahl-Hirschman Index", in *Paolo Baffi Centre Research Paper*, núm. 2009-55.

MARTÍNEZ-FAURA, Ú. y J. GÓMEZ-GARCÍA, 2001, "Modelos Migratorios: una revisión", en *Revista Asturiana de Economía*, núm. 21, pp. 209-235.

METZGER, W., 2006, *Laws of seeing*, The MIT Press, Cambridge, Mass. USA.

MODICA, G. y L. POGGIOLINI, 2013, *A first course in probability and Markov Chains*, John Wiley & Sons. Chichester, West Sussex, Inglaterra.

MULLIGAN, G.F., 1991, Equality measures and facility location, in *Papers in Regional Science*, vol. 70, núm. 4, pp. 345-365.

O'KELLY, M.E. y M.W. HORNER, 2003, "Aggregate accessibility to population at the county level: US 1940-2000", in *Journal of Geographical Systems*, vol. 5, núm. 1, pp. 5-23.

PARTIDA, V., 1989, "Aplicación de cadenas de Markov para proyecciones demográficas en áreas geopolíticas menores", en *Estudios Demográficos y Urbanos*, vol. 4, núm. 3, pp. 549-571.

PLANE, D.A. y G.F. Mulligan, 1997, "Measuring spatial focusing in a migration system", in *Demography*, vol. 34, núm. 2, pp. 251-262.

RABINER, L., 1989, "A tutorial on hidden Markov models and selected applications in speech recognition", in *Proceedings of the IEEE*, vol. 77, núm. 2, pp. 257-286.

RABINER, L., y B.H. JUANG, 1986, "An introduction to hidden Markov models", in *ASSP Magazine IEEE*, vol. 3, núm. 1, pp. 4-16.

RAPPAPORT, S., 1995, "Economic models and historical explanation", in *Philosophy of the Social Sciences*, vol. 25, núm. 4, pp. 421-441.

RASKALL, P. y G. MATHESON, 1992, Understanding the Gini Coefficient, in *Newsletter: Social Policy Research Center*, Universidad de Nueva Gales del Sur, vol. 6, Sidney, Australia.

ROGERS, Andrei y Stuart SWEENEY, 1998, "Measuring the spatial focus of migration patterns", in *The Professional Geographer*, vol. 50, núm. 2, pp. 232-242.

ROGERSON, P.A., 1979, "Prediction: a modified Markov Chain approach", in *Journal of Regional Science*, vol. 19, núm. 4, pp. 469-478.

ROMO-VIRAMONTES, R., Y. TÉLLEZ-VÁZQUEZ y J. LÓPEZ-RAMÍREZ, 2012, "Tendencias de la migración interna en México en el periodo reciente", en Conapo, *La situación demográfica de México 2013*, Conapo, pp. 83-106. México.

SINGH, S., 2001, "Multiple forecasting using local approximation", in *Pattern recognition*, vol. 34, núm. 2, pp. 443-455.

SPICKER, P.; S. LEGUIZAMON y D. GORDON, 2007, *Poverty: an international glossary*, Zed Books, NY, pp. 87.

SYS, C., 2009, "Is the container liner shipping industry an oligopoly?", in *Transport Policy*, vol. 16, núm. 5, pp. 259-270.

STARTZ, R., 2008, "Binomial autoregressive moving average models with an application to US recessions", in *Journal of Business & Economic Statistics*, vol. 26, núm 1, pp. 1-8.

VANDEWALLE, N., M. AUSLOOS y Ph. BOVEROUX, 1999, "The moving averages demystified", in *Physica A: Statistical Mechanics and its Applications*, vol. 269, núm. 1, pp. 170-176.

WATKINS, J.F., 1986, *Migration of the elderly in the United States: a multiregional analysis*, Tesis de Doctorado, University of Colorado, Boulder, Colorado, USA.

WEBBER, M.J., 1984, *Explanation, prediction and planning: the Lowry model*, Taylor & Francis. Londres.

YIN, G.G. y Q. ZHANG, 2010, *Continuous-Time Markov Chains and applications: a two-time-scale approach*, Springer. NY.

## INFORMACIÓN SOBRE LOS AUTORES

### *Carlos Garrocho*

Maestro en Desarrollo Urbano por El Colegio de México (1987) y Doctor en Geografía Socioeconómica por la Universidad de Éxeter (Inglaterra: 1992). Investigador Nacional Nivel 3. Integrante de la Comisión de Evaluación de Ciencias Sociales del Sistema Nacional de Investigadores (2014-2016). En El Colegio Mexiquense: Investigador desde 1986; Integrante de la Junta de Gobierno (2016); Miembro del Consejo Académico (desde 2008); Coordinador Académico (1994-1998); Fundador de la Revista *Economía, Sociedad y Territorio* (1994) y Director de la Revista (1994-1998 y nuevamente desde 2015). Ha publicado 22 libros científicos y 84 capítulos y artículos científicos nacionales e internacionales. Conferencista Magistral por parte de la Delegación Oficial Mexicana en las Cumbres de Población de 2014 y 2016 en la Sede de Naciones Unidas en Nueva York. Nombrado por el United Nations Population Fund: Coordinador de un grupo de expertos en el tema de ciudades sostenibles (desde 2014). Designado por la ONU como experto internacional en ciudades y representante de México en

la XXIII Cumbre Iberoamericana de Panamá (2013). Ganador del Premio Estatal de Ciencia del Estado de México (2011). Secretario de Planeación del Desarrollo del Gobierno del Estado de SLP (1997-2003). Su libro más reciente es: *Segregación socioespacial de la población mayor en la Ciudad de México: la dimensión desconocida del envejecimiento*, El Colegio Mexiquense, CDMX, México. En colaboración con Juan Campos. Dirección electrónica: cfgarrocho@gmail.com

*Eduardo Jiménez López*

Doctor en Ciencias Aplicadas por la Universidad Autónoma de San Luis Potosí. Es miembro del Sistema Nacional de Investigadores (Candidato desde 2014). Profesor-investigador de El Colegio Mexiquense A.C. Su publicación más reciente es “Estructura profunda de los flujos migratorios en México, 1990-2010”, en Conapo, *La situación demográfica de México 2014*, Conapo, México, en colaboración con Carlos Garrocho y José Antonio Álvarez Lobato. Su principal línea de investigación es la modelización de la migración interna en México. Dirección electrónica: eduardo.jimenez@ipicyt.edu.mx

*José Antonio Álvarez Lobato*

Doctor en Geografía, Maestro en Ingeniería e Ingeniero en computación. Miembro del Sistema Nacional de Investigadores. Secretario General de El Colegio Mexiquense, A.C., Coordinador del Laboratorio de Análisis Socioespacial, y miembro del Seminario de Estudios Estratégicos del Estado de México de la misma institución. Miembro de Comité Técnico Académico de la Red Temática Conacyt de Pobreza y Desarrollo Urbano hasta el 2015. Autor de diversos capítulos de libro y artículos científicos en temas de observación urbana, análisis espacial y economía urbana así como diversos artículos de divulgación en temas de TICs. Entre sus publicaciones se encuentran *La dimensión espacial de la competencia comercial; Observatorios urbanos en México: lecciones, propuestas y desafíos*; “Acceso a oportunidades: el principal desafío” en Brambila, Carlos (Ed.) *Prioridades de Investigación sobre Pobreza y Desarrollo Territorial, México: 2020*, ITESM, 2015; “Spatial organization of banking branches in the intra-metropolitan space in Toluca, Mexico” (coautoría con Carlos Garrocho) en Garrocho, C. (ed.). *Advances in commercial geography*, Mexico, El Colegio Mexiquense, 2013. “Calculation intraurban agglomeration of economic units with planar and network K-functions: a comparative analysis, en coautoría con Carlos Garrocho y Tania Chávez, en *Urban Geography*,

34(2), abril, 2013. Sus áreas de interés se centran en geografía urbana, análisis espacial, accesibilidad y movilidad urbana, sistemas de información geográfica en línea, y desarrollo de indicadores para el monitoreo urbano. Dirección electrónica: [jalvar@cmq.edu.mx](mailto:jalvar@cmq.edu.mx)

Artículo recibido el 8 de mayo de 2015 y aprobado el 1 de febrero de 2016.