

PATRONES DE ELECCIÓN DE LOS PARTICIPANTES DEL CONCURSO PARA EL INGRESO A LA EDUCACIÓN MEDIA SUPERIOR DE LA COMIPEMS

INGRID GARCÍA PINZÓN

Resumen:

Este estudio identifica si los jóvenes que concursan para ingresar al nivel medio superior solicitan escuelas de acuerdo con cierto patrón y establece si en ello influyen algunas variables. Se identificaron cuatro patrones de elección según el subsistema de las escuelas y se conformó un grupo por cada uno. Los datos se obtuvieron de un cuestionario aplicado en 2012 a 150 mil 404 jóvenes y se hizo una regresión logística multinomial. Las variables fueron: nivel socioeconómico, distancia entre los planteles elegidos y el hogar, promedio de secundaria y proporción de escuelas en que, según el puntaje obtenido en la prueba, el joven hubiera ingresado entre el número de las que solicitó. Esta última variable en el grupo que eligió como primera opción una escuela con un subsistema distinto al de la segunda no fue significativa; el resto sí lo fue para todos los grupos. Personas con distintas características tienen patrones de elección diferentes, lo cual puede indicar que ciertos grupos menos favorecidos no cuentan con información suficiente para elegir escuelas; se sugiere diseñar estrategias para ampliarla.

Abstract:

This study determines if the young people who apply for admission into high school select schools according to a certain pattern, and if variables influence the process. Four selection patterns were identified, according to the subsystem of schools; one group was formed for each pattern. The data were obtained from a questionnaire answered in 2012 by 150,404 young people; multinomial logistic regression was used for analysis. The variables were socioeconomic level, distance between the selected schools and the home, middle school average, and proportion of the requested schools that the young person could have entered based on the individual test results. This final variable was not significant in the group whose first choice was a school in a subsystem different from the second choice; the remaining variables were significant for all groups. Since individuals with different characteristics have different selection patterns, the possible indication is that certain disadvantaged groups may not have sufficient information to select schools; the suggestion is to design strategies to provide needed information.

Palabras clave: selección de estudiantes, educación media superior, nivel socioeconómico, desempeño académico, México.

Keywords: student selection, high school education, socioeconomic level, academic performance, Mexico.

Ingrid García Pinzón: estudiante de doctorado en la Universidad Nacional Autónoma de México-Facultad de Psicología. Avenida Universidad 3004, Ciudad Universitaria, 04510, México, Distrito Federal. CE: gpinzoni@gmail.com

Introducción

Numerosas investigaciones han tratado el tema de la elección de escuelas y han encontrado consistentemente que ya sean los padres o los posibles estudiantes con nivel socioeconómico bajo tienen una tendencia a elegir escuelas distintas de aquellos con un nivel alto (Montgomery, 2002; Zietz y Joshi, 2005). Estos hallazgos son relevantes porque, desde un punto de vista de la teoría de elección, el sistema que plantea a los posibles estudiantes las reglas por las cuales es asignado, es un “arquitecto de decisiones” (Thaler y Sunstein, 2008); las normas que establece pueden contribuir a que un individuo tenga preferencias sesgadas y, en el peor de los casos, provocar que ciertos grupos se vean desfavorecidos. Sin embargo, también es posible establecer un sistema cuyas reglas permitan a los jóvenes acceder a las que consideren las mejores escuelas y así consigan sus metas educativas (Thaler y Sunstein, 2008).

Diversos estudios que involucran el diseño de un sistema de asignación de alumnos a escuelas de nivel medio superior consideran una serie de reglas y restricciones e indican que las fallas en la elección de las opciones escolares determinan la insatisfacción y posible deserción de los estudiantes (Abdulkadiroglu, Pathak, Roth y Sonmez, 2005; Abdulkadiroglu y Sonmez, 2003).

Uno de estos sistemas es el concurso de selección para el ingreso al nivel medio superior organizado por la Comisión Metropolitana de Instituciones Públicas de Educación Media Superior (COMIPEMS). En este proceso participan los jóvenes (alrededor de 15 años de edad) que aspiran a un lugar en dicho nivel educativo de la Zona Metropolitana de la Ciudad de México que, para efectos de este concurso, se ha definido convencionalmente como la región que incluye el Distrito Federal y 22 municipios conurbados del Estado de México y que representa cerca de 20% de la población nacional (INEGI, 2010).

Al final de este proceso, los jóvenes continúan con su vida académica y el éxito o fracaso está relacionado con las características de los estudiantes y de las escuelas.

Este concurso pretende, fundamentalmente, resolver el problema de asignar a un grupo de jóvenes a alguna escuela para que todos los aspirantes que obtengan más de 30 aciertos en la prueba puedan ingresar a alguna institución y, a su vez, ninguna se quede sin estudiantes. Para lograrlo, cada uno de los 300 mil aspirantes deben elegir de una a veinte opciones

escolares y ordenarlas de tal manera que indiquen en primer lugar aquella que más prefieran, posteriormente la que desearían en segundo lugar y así sucesivamente. Posteriormente, los jóvenes presentan un examen de conocimientos constituido por 128 reactivos de opción múltiple, organizados en ocho campos temáticos (español, historia, geografía, formación cívica y ética, matemáticas, física, química, y biología) y que evalúan habilidades verbales y matemáticas básicas, así como habilidades y conocimientos específicos que se adquieren de manera regular en la escuela secundaria. Los estudiantes pueden obtener de cero a 128 aciertos. Para tener derecho a ingresar a una escuela en el 2012, los jóvenes debían contar con al menos 31 respuestas correctas.

Para asignarles escuela, todos los participantes son ordenados del que alcanzó el mayor número de aciertos al que obtuvo el menor. En un primer momento se elige al grupo con el puntaje más alto en la prueba y se le asigna la escuela elegida en primer lugar, posteriormente se revisa a los del segundo mejor puntaje y, en caso de que aún exista lugar en la primera opción, se les sitúa en ésta y así sucesivamente.

Los participantes empatados en cuanto a su número de aciertos y que hayan elegido la misma escuela son asignados a ella siempre y cuando existan lugares suficientes; en caso contrario, ese grupo deberá conformarse con su segunda opción. Esto ocurriría si el empate fuera en la primera escuela elegida, pero si se tratara de la segunda, los participantes tendrían que conformarse con su tercera opción y así sucesivamente.

Para los aspirantes es una tarea complicada, sobre todo si no es claro cuál es la escuela que más les conviene y cuál es la información relevante para su elección. Además, esta situación es nueva para ellos y es probable que no les sean claras cuáles son las restricciones y reglas que les plantea el concurso de ingreso. Por otro lado, elegir veinte opciones de entre aproximadamente 600 posibles, ordenarlas jerárquicamente y, que esto dependa del desempeño académico que muestren en la prueba de logro, hace aún más complejo el proceso, ya que existe la posibilidad de que los asignen a escuelas que no son de su preferencia, que se encuentren lejanas a su hogar o que las consideren de mala calidad.

Quizá haya aspirantes que elijan de acuerdo con reglas complejas y otros simplemente lo hacen de manera natural; es decir, sin considerar rebuscamientos o factores no relacionados con la situación. Se ha encontrado que bajo ciertas circunstancias, quienes eligen de manera sencilla pueden

obtener mejores resultados que aquellos que lo hacen de modo complicado (Apestequia y Ballester, 2011).

Thaler y Sunstein (2008) encontraron que debido a la gran cantidad de información a la que se enfrentan los concursantes es muy posible que utilicen “pistas” que los lleven a elegir opciones que les parecen las más “recomendadas”. Por otro lado, debido a que primero hacen la elección y posteriormente el examen, se puede generar la sensación de que lo que elijan o la manera en que lo hacen puede tener consecuencias mayores de las que realmente tiene. Así, el interés de este estudio se basa en conocer si los aspirantes utilizan alguna estrategia para ordenar las escuelas de su preferencia. Algunas de estas “pistas” pueden ser las siguientes:

- 1) El subsistema al que pertenece cada escuela.
- 2) El resultado que los concursantes consideran que obtendrán en el examen con respecto al resto de los participantes.
- 3) La distancia entre las escuelas y el hogar de los aspirantes.
- 4) Promedio obtenido en la secundaria.
- 5) El nivel socioeconómico de los concursantes.

Respecto del subsistema, O’Shaughnessy (2007) encontró que los padres de familia reportaban constantemente que la calidad de la escuela fue una variable importante para ellos, ya que consideraron que algunos subsistemas son de mejor calidad que otros por las características de sus alumnos, lo que tiene un efecto sobre la selección de escuelas para sus hijos. Brasington y Hite (2012) reportaron que las opiniones sobre la calidad escolar son tan importantes como las medidas objetivas, por ejemplo, el resultado de las escuelas en evaluaciones por medio de pruebas estandarizadas.

Sin embargo, para un joven recién egresado de la secundaria y para su familia puede ser muy complicado identificar qué escuela es de mejor calidad que otra, sobre todo si no existe una evaluación estandarizada que incluya a todas las involucradas en el concurso. Por ejemplo, la Evaluación Nacional del Logro Académico en Centros Escolares (ENLACE) que se ha aplicado en el nivel medio superior ofrece información específica sobre el desempeño académico dentro de las escuelas; sin embargo, no todas las instituciones del nivel medio superior realizan esta prueba. Debido a esta falta de información, probablemente los jóvenes y sus familias consideren al subsistema como uno de los indicadores más importantes de calidad

escolar. Evidentemente, las escuelas que pertenecen a ciertos subsistemas son mucho más grandes y cuentan con mayor presupuesto que el resto.

Sobre el resultado que los concursantes consideran que obtendrán en el examen con respecto al resto de los participantes (inciso 2), destaca que el cálculo de la oportunidad que tiene un estudiante para ingresar a determinada escuela es independiente del ordenamiento que hace de sus preferencias, porque esta clasificación determinará en qué momento el sustentante tiene la oportunidad de empezar a competir en el concurso. Es decir, si elige como primera opción las más difícil o con mayor demanda (esto último también se relaciona con la dificultad, ya que entre más solicitada sea una escuela, más difícil es ingresar a ella) podrá comenzar a competir desde el principio; sin embargo, si su primera opción es la escuela menos demandada o la más fácil de ingresar comenzará a competir por un lugar en ella hasta el final. Pero la probabilidad de ser aceptado por una escuela está relacionada con el puntaje que se obtenga en la prueba y las opciones que se proponen.

De acuerdo con lo anterior, es evidente que no solo se requiere una cuidadosa revisión de los subsistemas a los que pertenecen las instituciones involucradas en el concurso, sino también considerar cuál será el desempeño de cada participante con respecto a los demás para poder identificar un patrón de elección. Ésta es quizá la tarea más difícil a la que se enfrentan los concursantes, ya que deben calcular cuál será su desempeño y el del resto de los participantes. Parece imposible; sin embargo, la teoría de juegos puede ser útil para hacer frente a las complicaciones que este concurso presenta, ya que estudia las “reglas del juego” (Roth, 2002). Es decir, los estudiantes pueden seleccionar opciones que no sean de su entera satisfacción, pero al saber que su nivel de desempeño en el examen será bajo pueden, por ejemplo, considerar una escuela que no sea de las más demandadas (de esta forma competirán con estudiantes de más bajo rendimiento académico) y podrían asegurar un lugar (Sonmez, 1997).

Otra manera en la que podrían actuar es que, como lo mencionó Cantalá (2010), los participantes no reporten sus “verdaderas preferencias”; sin embargo este sistema, tal como está diseñado, no propicia que esto ocurra, sobre todo porque si un participante coloca en primer lugar una escuela B donde sea más fácil ingresar y en segundo lugar una A, difícil, y obtiene un puntaje tan alto como para ingresar a dicha escuela, lo que ocurrirá es que ese aspirante entrará a la B. Para que un concursante se

incorpore a una escuela muy demandada debe conseguir un puntaje alto, sin importar el orden en que reporte sus preferencias.

Se puede pensar que los estudiantes se autoseleccionan, es decir, aquellos con un mejor desempeño académico solicitan las mejores escuelas, este efecto se ve mermado en este concurso porque pueden elegir cualquier escuela sin ningún costo para ellos. Es decir, un aspirante con bajo rendimiento académico puede solicitar una escuela con alta demanda y no debe pagar ningún precio por hacerlo, de cualquier modo debe presentar el examen y tiene la oportunidad de ingresar a una buena escuela o, en su defecto, probablemente no la tiene, pero no pierde nada al solicitarla. Sin embargo, al no elegir como primera opción una institución en la que pueden ingresar dado su rendimiento académico, están desperdiciando una oportunidad, aunque ésta sea una de veinte.

Por otro lado, de acuerdo con la Encuesta Nacional de Deserción en la Educación Media Superior (SEP, 2012), 7.6% de los jóvenes que abandonaron los estudios en esta etapa dijeron haberlo hecho debido a “que la escuela les quedaba muy lejos”. Por lo tanto, qué tan lejos o cerca se vive de la escuela puede desempeñar un papel fundamental en la elección (inciso 3).

De acuerdo con Chumacero, Gómez y Paredes (2011) y Lang, Collins y Kearns (2011), los padres eligen las escuelas para sus hijos considerando la distancia como factor principal. Idealmente se espera que seleccionen escuelas cercanas a su domicilio, sin embargo, la calidad percibida de las mismas puede motivarlos a considerar las lejanas.

Por otra parte, algunos autores (Allen, 1999; Beecher y Fischer, 1999; McGrath y Braunstein, 1997) consideran el promedio general de calificaciones (inciso 4), como un indicador simple e importante del éxito escolar. Además, éste no únicamente refleja los conocimientos y habilidades adquiridas, sino también características propias del alumno, por ejemplo los recursos cognitivos y no cognitivos con los que cuenta y que son elementos de igual importancia para el desempeño académico, el ingreso y el éxito en la vida social (Heckman, Stixrud y Urzua 2006; Murnane, Willett y Levy, 1995). Así, este factor se incluyó como posible “pista” debido a que es una fuente de información con la que cuentan los jóvenes como referencia de su desempeño académico, lo cual es una de las variables más relevantes que determinarán su resultado en el concurso.

Respecto del nivel socioeconómico, algunos investigadores han destacado la importancia de las diferencias que éste provoca (inciso 5) o de los recursos

materiales de una familia en el trayecto de la vida de sus hijos (Duncan *et al.*, 1998; Huston, 1999). El nivel socioeconómico se constituye por el capital cultural (escolaridad de los padres y acceso a la información de los estudiantes) y el económico (bienes materiales de los padres) (Schulz, 2005). Estas variables se encuentran íntimamente relacionadas, ya que los padres transforman capital económico en cultural cuando deciden invertir en la adquisición de recursos que facilitan el aprendizaje y apoyan el desempeño académico de sus hijos (por ejemplo, libros o acceso a Internet) (Bourdieu, 1977; Fernández, Banegas y Blanco, 2004; Sullivan 2002).

Como se mencionó al inicio de este estudio, el nivel socioeconómico ha sido una variable clave cuando se trata de seleccionar una escuela; además, se sabe de antemano que ésta afecta el desempeño académico de manera sustancial. Estudios como el de Montgomery (2002) y Zietz y Joshi (2005) coinciden en que dicho nivel y la raza son características que desempeñan un papel central en la elección de escuelas.

Se puede argumentar que el nivel socioeconómico no es una variable considerada para la selección de estudiantes en el concurso de ingreso de COMIPEMS; sin embargo la evidencia provista por Duncan *et al.* (1998), Huston (1999), Montgomery (2002) y Zietz y Joshi (2005) sugiere que puede influir en dos momentos de este proceso de selección: *a)* cuando los aspirantes eligen las escuelas de su preferencia y *b)* cuando presentan el examen.

En este estudio se construyó un indicador de nivel socioeconómico, es importante señalar que la escala obtenida es intervalar, lo que proporciona enormes ventajas ya que este grado de medición permite una gran cantidad de operaciones estadísticas. El desarrollo de la escala se realizó a partir de los datos de los cuestionarios de contexto y con base en el modelo de Crédito Parcial. Éste pertenece, a su vez, a la familia de los modelos de la teoría de respuesta al ítem (Wright y Masters, 1982). La razón para utilizarlo fue que resulta el más apropiado para preguntas con distinto número de opciones de respuesta; además, fue diseñado especialmente para respuestas parcialmente correctas. El modelo fue desarrollado por Masters (1982) y Wright y Masters (1982) y comparte las características de la familia de modelos de Rasch.

El objetivo principal de este trabajo fue identificar patrones (basados en el subsistema) en la elección de los participantes en este concurso y dar cuenta de si el resultado que consideran que tendrán en el examen con

respecto al resto de los participantes, la distancia a la que se encuentran las escuelas de su hogar, su promedio de secundaria y su nivel socioeconómico contribuyen a que los jóvenes hayan reportado cada patrón identificado.

Método

Participantes

Se utilizó la base de datos de los estudiantes que participaron en el concurso de la COMIPEMS en 2012, que contenía toda la información requerida para realizar este estudio. La muestra fue de 150 mil 404 jóvenes, 56.0% hombres y 44.0% mujeres, con una media de edad de 15.7, con una desviación estándar de 1.8.

Instrumentos

Como instrumentos se utilizaron el examen de selección para medir el desempeño académico de los jóvenes y el cuestionario de contexto, aplicado en 2012, que contestan cuando se registran en el concurso, y de los que se obtuvo la base de datos aquí descrita. De ella se seleccionaron únicamente a los estudiantes que respondieron el cuestionario y registraron adecuadamente todos sus datos en el proceso de inscripción al concurso. Además se consultaron las bases de datos *Consulta Códigos Postales* (Correos de México, 2014) y *GeoNames* (2014) como fuentes para el cálculo de las distancias.

Procedimiento

Para el uso de la base de datos se le solicitó, por escrito, a la COMIPEMS su autorización. Después se procedió a eliminar de la base a las personas que no habían respondido el cuestionario de contexto o no habían reportado adecuadamente los datos de su registro (por ejemplo su código postal).

Posteriormente, se construyeron los indicadores que fungieron como variables independientes de este estudio. Para cada uno se realizó lo siguiente:

- 1) *Patrón de acuerdo con el subsistema al que pertenece cada escuela.* Se clasificó cada una de las elecciones de los 150 mil 404 aspirantes de acuerdo con subsistema de cada escuela elegida. Posteriormente se revisaron todos los patrones de elección por medio de la frecuencia con la que cada joven eligió una escuela con el mismo subsistema. El objetivo en este caso fue revisar si un participante elegía muchas veces escuelas con el mismo subsistema.

- 2) Proporción obtenida de la división del número de escuelas en las que se podía quedar el aspirante entre el número de escuelas solicitadas. Esta variable se construyó calculando el puntaje más bajo con el que el último participante ingresó a cada escuela (punto de corte), posteriormente se calculó si el puntaje de cada joven era mayor o igual a cada punto de corte. Finalmente, se dividió el número de escuelas en las que el participante “hubiera ingresado”¹ con el puntaje que obtuvo entre el número de escuelas que solicitó. Al obtener el número 1 como resultado de esta división, indica que ese joven tenía un puntaje en el examen suficiente para quedarse en cualquiera de las escuelas que eligió; 0.5, que obtuvo un puntaje con el cual se hubiera quedado en la mitad de los establecimientos escogidos; y finalmente el 0 indica que el joven tenía un puntaje muy bajo como para quedarse en las que solicitó. Este indicador da cuenta de la proporción del número de escuelas que eligió cada aspirante y el número de aquellas en las que se hubiera quedado.
- 3) *Distancia entre las escuelas elegidas y el hogar de los aspirantes.* Utilizando los códigos postales, se calculó la distancia entre el domicilio del participante y cada una de las escuelas que seleccionó. Al código se le asoció con sus coordenadas x y y , estos datos se obtuvieron por medio de un listado de la página de Correos de México (2014) y fue cotejada con las coordenadas x - y extraídas de *GeoNames* (2014). Posteriormente, con el teorema de Pitágoras, se calculó la distancia lineal entre los dos puntos, es decir, entre la ubicación del código postal del domicilio de los participantes y la de las escuelas seleccionadas, finalmente se calculó un promedio de distancia de todas las opciones.
- 4) *Promedio de secundaria.* Esta información se extrajo de la base de datos, que es el promedio reportado por los jóvenes en el momento de solicitar su participación en el concurso.
- 5) *Nivel socioeconómico de los concursantes.* Se construyó una escala de nivel socioeconómico considerando las siguientes características: escolaridad de la madre y del padre, así como el número tanto de libros como de computadoras que los jóvenes poseen en su hogar. Se utilizó el modelo de Crédito Parcial de Masters (Wright y Masters, 1982), que se establece a partir de la determinación de la probabilidad condicional de responder a una de las opciones de respuesta o

categorías de una serie de preguntas politómicas. El análisis se llevó a cabo de la siguiente manera:

En un primer momento, se calcularon los valores óptimos de ajuste al modelo para cada pregunta que pretendía medir nivel socioeconómico, estos valores son: para el OUTFIT e INFIT de 0.7 a 1.3 (Gustafsson, 1980). INFIT es un estadígrafo sensible a respuestas no esperadas a reactivos calibrados cerca del nivel de habilidad del sustentante, mientras que el OUTFIT es también sensible a respuestas no esperadas pero lejos de las habilidades calibradas para los sustentantes o para las dificultades de los reactivos (Gustafsson, 1980). A continuación se procedió a calcular la varianza explicada por un único factor, los valores deben ser mayores a 60% (Hair *et al.*, 1998 y 1999 y Pett, Lackey y Sullivan, 2003). Posteriormente se obtuvo el alfa de Cronbach (consistencia interna), cuyos valores deben ser mayores al 0.70 (Nunnally y Bernstein 1994; DeVellis, 2003). Finalmente, se construyó el puntaje de la escala, el cual fue generado en *logits*.

Una vez obtenidos los indicadores enlistados en los incisos anteriores, se realizaron correlaciones entre las variables que se consideraron posibles “pistas” para la elección de escuelas con el objeto de mostrar que no existía colinealidad entre las variables independientes. Después se elaboró un modelo de regresión logística multinomial y se obtuvieron las probabilidades de que una persona tuviera un patrón de elección en específico de acuerdo con el subsistema según el resto de las variables aquí mencionadas. La regresión logística multinomial es útil para los casos en los que se desea predecir que una persona con ciertas características (variables independientes) pertenezca a un grupo en específico (pueden haber tres o más grupos). Es similar a un modelo de regresión lineal pero está adaptado para modelos en los que la variable dependiente es politómica, como en este caso donde la variable dependiente son los grupos que se conforman por los patrones de elección que se identificaron en este estudio. Los coeficientes de regresión logística pueden utilizarse para estimar qué factores contribuyen a la probabilidad de pertenecer a cierto grupo de cada variable independiente del modelo, tomándose como referencia un grupo en particular. La regresión logística no se basa en supuestos distribucionales en el mismo sentido en que lo hace el análisis discriminante. Adicionalmente, al igual que con otras formas de regresión, la multicolinealidad entre los predictores puede llevar a

estimaciones sesgadas y a errores estándar inflados, por esta razón, primero se estimaron las correlaciones entre variables independientes.

La regresión logística multinomial es más eficaz cuando la pertenencia a grupos es una variable categórica auténtica, como es el caso de este estudio: grupos conformados por el patrón de elección que realizan.

Resultados

A continuación se presentan los resultados de cada una de las variables que se incluyeron en el modelo. Posteriormente, se enuncian las correlaciones entre ellas. Finalmente, se describen los valores del modelo de regresión logística multinomial junto con las probabilidades de que una persona tenga un patrón de elección específico de acuerdo con el subsistema.

Patrón de acuerdo con el subsistema al que pertenece cada escuela

En cuanto a la elección en este concurso, 55.2% de los aspirantes eligieron al menos la primera y la segunda opciones (pudieron ser las tres, cuatro o cinco primeras) de escuelas con el mismo subsistema. De este subconjunto, 18.2% reporta exactamente el mismo exclusivamente en su primera y segunda opciones. Suponiendo que para los estudiantes las primeras escuelas que eligen son mejores que las últimas, parece que consideran el subsistema como algo muy importante. Evidentemente no hay datos suficientes para afirmar esto; sin embargo, no deja de ser interesante que los patrones identificados que se reportan en estos resultados únicamente se realizaron considerando los subsistemas de las opciones educativas. En los porcentajes reportados en el cuadro 1 se incluyeron todos los subsistemas. Además, es importante resaltar que alrededor de 50% de los sustentantes indicaron un plantel de la Universidad Nacional Autónoma de México (UNAM), que consta de 14 opciones, a pesar de que en total hay alrededor de 600 disponibles. Independientemente de esto, se encontró que aproximadamente 50% de los jóvenes escogió este subsistema en primer lugar, sin importar el patrón de elección que fue identificado.

El cuadro 1 describe los cuatro tipos de patrones de respuesta que se encontraron de acuerdo con el tipo de subsistema al que pertenecían las escuelas elegidas.

A partir de este punto se dividió a la población en cuatro grupos de acuerdo con los patrones reportados y explicados en el cuadro 1, se les llamará a estos grupos: “Sin patrón”, “primeras iguales”, “últimas iguales”

y “alternadas”. Esta será la variable dependiente que se refiere al tipo de patrón que considera el subsistema de las opciones educativas que eligieron los jóvenes.

CUADRO 1

*Patrones de elección de acuerdo con subsistema de las escuelas
(se deben considerar para cada caso A, B y C como subsistemas distintos)*

Patrón de elección	Características del patrón de elección	% del total de N
Sin patrón	No se logró encontrar algún patrón entre las opciones que eligieron	4.3
Primeras iguales	Eligieron al menos la primera y la segunda opciones con el mismo subsistema y el resto pertenecía a un subsistema distinto. También pudieron elegir las tres, cuatro o cinco primeras escuelas con el mismo subsistema y posteriormente, de igual forma, las siguientes opciones incluidas eran de un subsistema diferente al de las primeras. Ejemplo: AAABBBBC...	55.2
Últimas iguales	Eligieron al menos las primeras tres opciones con un subsistema distinto pero, a partir de la cuarta, reportaron el mismo subsistema en al menos dos opciones (las cuales ocuparían el lugar tres y cuatro). En términos generales, eligieron las primeras opciones con subsistemas variados, pero las que reportaron al final eran del mismo subsistema. Ejemplo: ABCCCC	34.2
Alternadas	Eligieron en primer y tercer lugares una escuela perteneciente a un subsistema distinto a la que consideraron en segundo lugar. Este patrón se puede repetir. Alternaron sus preferencias entre dos subsistemas al principio de sus elecciones reportadas, colocaban primero una opción correspondiente a un subsistema, en segundo, una con subsistema diferente a la primera opción, pero en tercer lugar volvieron a reportar una escuela con el mismo subsistema a la primera. Ejemplo: ABABCA...	6.3
Total		100.0

En cuanto a los resultados de las variables independientes, en primer lugar se presenta la proporción (división) del número de escuelas que solicitó el joven entre el de aquellas en las que se hubiera quedado según su puntaje en el examen. Debido a que esta variable fue medida con una división, el número máximo a obtener es 1, lo que indicaría que ese joven eligió, por ejemplo, tres escuelas y en las tres se hubiera quedado. Si este mismo joven eligiera tres escuelas pero, dado el puntaje que obtuvo, se hubiera quedado únicamente en dos de ellas, entonces esta proporción sería igual a 0.66 o $2/3$. Los estadísticos descriptivos de esta variable fueron: una media de 0.56 y una desviación estándar de 0.35.

Sobre la distancia lineal a la que se encuentra cada una de las escuelas, obtenida por medio del teorema de Pitágoras, los estadísticos descriptivos fueron: una media de 0.27 y una desviación estándar de 0.28, considerando las coordenadas x-y de cada código postal que contenía la base de datos de Correos de México (2014) y cotejándola con la base de *GeoNames* (2014).

Los estadísticos descriptivos de la variable de promedio de secundaria obtenida de los datos del cuestionario de contexto fueron: media de 7.27 y desviación estándar de 2.49.

Para calcular el nivel socioeconómico de los concursantes se utilizó el procedimiento para elaborar una escala con la Teoría de Respuesta al Ítem (TRI) mediante el modelo de Crédito Parcial (Wright y Masters, 1982). Las principales características de las escalas obtenidas fueron:

- 1) Las escalas son unidimensionales, es decir, cada escala mide un mismo constructo y por lo tanto se componen de reactivos correlacionados entre sí.
- 2) Siguiendo el principio de parsimonia, cada reactivo debe aportar información sustancialmente para su inclusión en la escala.
- 3) Los reactivos presentan valores de INFIT y OUFIT dentro del rango establecido (.7 a 1.3) (Gustafsson, 1980).

Los puntajes obtenidos para cada escala tienen como unidad los lógitos, con una media teórica de cero y desviación estándar de uno. Como criterios para el funcionamiento de una escala como indicador es necesario que se explique al menos 60% de la varianza (Hair *et al.*, 1998 y 1999 y Pett, Lackey y Sullivan 2003) y un coeficiente de confiabilidad (alfa de Chronbach) de 0.70 como mínimo (Nunnally y Bernstein, 1994; DeVellis, 2003).

Se incluyeron las siguientes variables para la escala de nivel socioeconómico: “cuántos libros hay en tu casa”, “escolaridad del padre”, “escolaridad de la madre”, y “si tienes computadora”.

CUADRO 2

Características psicométricas de la escala de nivel socioeconómico

VARIABLES	INFIT	OUTFIT	Correlación con la escala
Número de libros	0.93	1.12	0.73
Escolaridad del padre	0.96	1.01	0.80
Escolaridad de la madre	0.80	0.85	0.78
Número de computadoras en el hogar	0.79	0.76	0.75

La escala tuvo un alfa de Cronbach de 0.66, que pudiera considerarse baja; sin embargo no lo es, dado que solo cuenta con cuatro reactivos. Cuando la cantidad de reactivos es muy baja, pueden darse valores más bajos de los tradicionalmente esperados y esto aún indica que la escala tiene una buena consistencia interna (Cortina, 1993; Nunnally y Bernstein 1994; Streiner, 2003). Los estadísticos descriptivos de esta variable, que se incluirá más adelante en un modelo como variable independiente, fueron: media de nivel socioeconómico igual a -2.54 lógitos y la desviación estándar igual a 2.56 lógitos.

En el cuadro 3 se enuncian las correlaciones correspondientes a las variables que posteriormente se incluirán en el modelo de regresión logística multinomial como variables independientes.

En general, las correlaciones son bajas, lo cual es adecuado para correr una regresión logística multinomial, ya que indica que no hay colinealidad ni redundancia entre las variables independientes. Dado lo anterior, se planteó el siguiente modelo:

La probabilidad de que $Y=1$, es decir de que un joven reporte un patrón de “primeras iguales” sea 1, en presencia de las variables independientes: x_1 =distancia,

x_2 =nivel socioeconómico, x_3 = proporción del número de escuelas en las que se podía quedar/el número de escuelas solicitadas y x_4 = promedio de secundaria.

CUADRO 3
Correlaciones entre las variables que se incluyeron en el modelo de regresión logística multinomial

	Promedio de secundaria	Proporción de núm. de escuelas en que se podía quedar/núm. de solicitadas	Distancia	Nivel socioeconómico
Promedio de secundaria				
Proporción de número de escuelas en que se podía quedar/número de escuelas solicitadas	.214**			
Distancia	.030**	.068**		
Nivel socioeconómico	.040**	.082**	.014**	

** La correlación es significativa al nivel 0.01 (bilateral).

Hay cuatro categorías: sin patrón, primeras iguales (categoría de referencia), últimas iguales y alternadas. Por lo tanto se modelan tres lógitos simultáneamente:

- (sin patrón/ primeras iguales | X) = $a_1 + b_1x$
- (últimas iguales / primeras iguales | X) = $a_2 + b_2x$
- (alternadas / primeras iguales | X) = $a_3 + b_3x$

La covariable X es común en cada lógito, pero se estiman coeficientes b_i diferentes para cada uno (incluso diferente constante, a_i). Consideremos una variable de respuesta polinómica Y con cuatro categorías de respuesta que denotaremos por Y_1, Y_2, Y_3 y Y_4 . Se pretende explicar la probabilidad de cada categoría de respuesta en función de un conjunto de covariables observadas $X = x_1, x_2, x_3, x_4$.

En los cuadros siguientes se presentan los resultados del modelo de regresión logística multinomial:

CUADRO 4
Contrastes de la razón de verosimilitud

Efecto	Criterio de ajuste del modelo -2 log verosimilitud del modelo reducido	Contrastes de la razón de verosimilitud		
		Chi-cuadrado	gl	Sig.
Intersección	293202.363	9599.001	4	0.000
Distancia	283837.792	234.429	4	.000
Nivel socioeconómico	284097.757	494.395	4	.000
Proporción de escuelas en las que se podía quedar/escuelas solicitadas	283715.253	111.891	4	.000
Promedio de secundaria	283881.249	277.886	4	.000

El estadístico de chi-cuadrado es la diferencia en las -2 log verosimilitudes entre el modelo final y el reducido. Este último se forma omitiendo un efecto del modelo final. La hipótesis nula es que todos los parámetros de ese efecto son cero. Los efectos de todas las variables tuvieron parámetros significativos. La significación estadística indica que el modelo con la nueva variable introducida mejora el ajuste de forma importante con respecto a lo que se tenía.

CUADRO 5
Pseudo R-cuadrado

Cox y Snell	.083
Nagelkerke	.095

El R cuadrado de Cox y Snell se basa en la comparación del log de la verosimilitud para el modelo respecto del log de la verosimilitud para un modelo de línea base. En este caso, el valor indica que 8.3% de la variación de la variable dependiente (patrón de elección de acuerdo con subsistema) es explicada por las variables independientes (distancia, nivel socioeconómico, proporción

de escuelas en las que se podía quedar/escuelas solicitadas y promedio de secundaria). El R cuadrado de Nagelkerke corrige la escala para cubrir el rango completo de 0 a 1, a diferencia del de Cox y Snell que tiene un valor máximo inferior a 1, con este R cuadrado se explica el 9.5% de la varianza del patrón de elección según el subsistema de las opciones reportadas de los jóvenes.

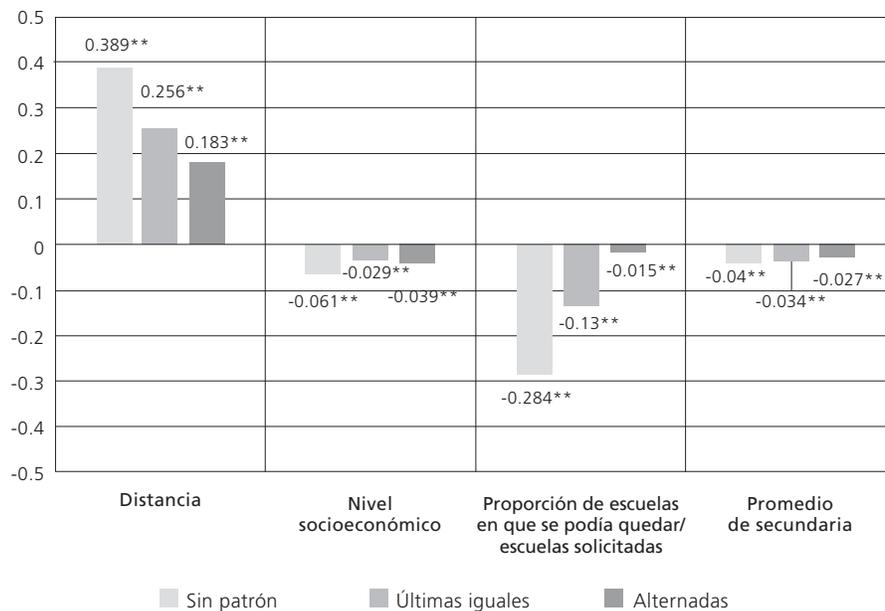
CUADRO 6
Regresión logística multinomial

Patrón de elección	B	Error est.	Wald	Gl	Sig.	Exp(B)	Intervalo de confianza a 95% para Exp(B)	
							Límite inf.	Límite sup.
Sin patrón	Intersección	-2.383	.042	3185.057	1	0.000		
	Distancia	.389	.044	78.831	1	.000	1.475	1.354 1.607
	Nivel socioeconómico	-.061	.005	159.456	1	.000	.941	.932 .950
	Proporción de escuelas en las que se podía quedar/escuelas solicitadas	-.284	.038	56.265	1	.000	.753	.699 .811
	Promedio de secundaria	-.040	.005	67.366	1	.000	.961	.952 .970
Últimas iguales	Intersección	-.304	.019	252.812	1	.000		
	Distancia	.256	.020	167.149	1	.000	1.291	1.242 1.342
	Nivel socioeconómico	-.029	.002	170.282	1	.000	.971	.967 .976
	Proporción de escuelas en que se podía quedar/escuelas solicitadas	-.130	.017	60.530	1	.000	.878	.850 .907
	Promedio de secundaria	-.034	.002	240.222	1	.000	.967	.962 .971
Alternadas	Intersección	-2.119	.037	3297.867	1	0.000		
	Distancia	.183	.038	23.409	1	.000	1.201	1.115 1.294
	Nivel socioeconómico	-.039	.004	88.235	1	.000	.962	.954 .970
	Proporción de escuelas en que se podía quedar/escuelas solicitadas	-.015	.032	.205	1	.651	.986	.925 1.050
	Promedio de secundaria	-.027	.004	39.794	1	.000	.974	.966 .982

Nota: La categoría de referencia es: Primeras iguales.

Como se puede observar en la gráfica 1, a mayor distancia, es menos probable pertenecer a la categoría de “primeras iguales”; sin embargo, no ocurre esto con las otras tres variables independientes incorporadas al modelo.

GRÁFICA 1
*Coefficientes del modelo de regresión
logística multinomial*



** . La B fue significativa al nivel 0.01 (bilateral).

Nota: La categoría de referencia es: Primeras iguales.

Entre mayor nivel socioeconómico, mayor promedio de secundaria, y entre más grande sea el número de escuelas que el joven eligió donde se podía quedar, mayor es la probabilidad de haber tenido un patrón de “primeras iguales”, que fue el patrón más frecuente.

Todas las variables fueron significativas, excepto en el caso de la categoría de “alternadas”, donde la proporción de escuelas en las que se podía

quedar el concursante entre las opciones solicitadas no parece contribuir significativamente a la probabilidad de que un aspirante elija este patrón sobre el patrón de “primeras iguales”.

Discusión

La mayoría de los participantes de este concurso elige con un patrón que se puede identificar con relativa facilidad y está basado en el subsistema de las escuelas, como se muestra en el cuadro 1, donde se presentan los tres patrones de elección identificados en este estudio. Esto muestra que las personas, cuando son enfrentadas a cantidades muy grandes de información, suelen reducirla tomando en cuenta algunas “pistas”, tal como lo afirman Thaler y Sunstein (2008).

De las variables independientes aquí consideradas, vale la pena destacar que con el modelo planteado solo se logró explicar 9.5% de la varianza de los grupos, lo cual indica que quizá no se estén incluyendo las variables más adecuadas. Sin embargo, se encontraron diferencias significativas en cuanto al patrón que eligen los sustentantes considerando la distancia; quienes consideran escuelas más lejanas son aquellos que no optan por un patrón o que eligen uno o los patrones de “últimas iguales” y “alternadas”. Éstos son más complejos en diseño con respecto a elegir las primeras opciones iguales; en general, estos sustentantes reportan patrones más rebuscados. Esto, de algún modo, también puede estar asociado con el nivel socioeconómico y por lo tanto al acceso que tienen los jóvenes a las escuelas. De esta manera, es posible que quienes prefieran escuelas lejanas a su domicilio sea porque consideran que las cercanas son de baja calidad y les generan menos oportunidades; si esto es cierto, la variabilidad en sus patrones de elección puede reflejar que no quieren una institución cercana a ellos, pero no tienen claro cuáles son las opciones que prefieren.

Por otro lado, el nivel socioeconómico es una variable que también contribuye para predecir el tipo de patrón con el que elegirán los jóvenes. Al parecer, los que tienen un nivel significativamente más bajo optan por patrones distintos al de “primeras iguales”. Esto parece deberse, al igual que en el caso de la variable anterior, a la falta de acceso a la información que enfrentan los grupos más desprotegidos. En este caso en particular, probablemente se deba que no están seguros o no tienen referencias de cuál es la mejor opción de escuela para ellos.

Los sustentantes que eligen escuelas donde no podrían quedarse tienden a optar por (utilizando como referencia el patrón de “las primeras opciones iguales”) “sin patrón” y “últimas iguales”. Finalmente, los que tienen menor promedio de secundaria escogen “sin patrón”, “últimas iguales” y “alternadas”. En el caso de estos dos últimos, la variabilidad de escuelas que aparecen en ellos puede deberse a los estudiantes que saben que tienen un bajo rendimiento académico también saben que les será muy difícil ingresar a una escuela de alta demanda. Sin embargo, dado que la estructura del concurso no castiga la solicitud de escuelas de alta demanda, este grupo de jóvenes de bajo rendimiento las escoge con la esperanza de poder ingresar; no obstante, como consideran que es poco probable, también incluyen escuelas donde saben que el ingreso es más accesible y por lo tanto sí tienen oportunidad de quedarse.

En términos generales, como se aprecia en la gráfica 1, parece que el patrón más común (“primeras iguales”) es elegido por los aspirantes con mayor nivel socioeconómico, por aquellos que solicitan escuelas donde pueden quedarse y consideran, en general, escuelas con menor distancia o más cercanas.

Este estudio muestra que los sustentantes más desprotegidos son aquellos que tienen un patrón de elección distinto al de “primeras iguales”, lo cual podría ser debido a la falta de información con la que cuentan.

Apesteagua y Ballester (2011) encontraron que bajo ciertas circunstancias, quienes no eligen de manera sofisticada pueden obtener mejores resultados que aquellos que lo hacen de manera muy complicada, este estudio aporta una posible razón de que esto ocurra, ya que parece que en el caso particular de este concurso, los jóvenes con menos recursos, que eligen en promedio las escuelas más lejanas y que consideran aquellas donde no se pueden quedar, dado su desempeño académico, son los que tienen patrones de elección más complicados. Es posible que esta complejidad esté asociada con el nivel socioeconómico y, en consecuencia, a las poblaciones más vulnerables, las de menos recursos económicos y las que tienen menor acceso a información relevante para su toma de decisiones.

Un sistema que permitiera a todos los sustentantes acceder a información más clara, que ayudara a los jóvenes y a sus familias a comprender las consecuencias de sus elecciones en su futuro académico, puede contribuir a que la mayoría de los aspirantes tengan las mismas oportunidades de elegir el conjunto de escuelas que sea más adecuada para ellos, sobre todo

para aquellos más desprotegidos, con nivel socioeconómico bajo o que solicitan escuelas lejanas a su domicilio.

El problema más grave de falta de información a la que se enfrentan estos concursantes es que no existe ningún indicador comparativo para escuelas de distintos subsistemas, lo que sí les puede ser muy evidente es que las escuelas son distintas porque pertenecen y son reguladas por diferentes instituciones. Además, no es claro si las escuelas de diversos subsistemas son comparables entre sí, dadas otras características como la ubicación o el tipo de educación que ofrecen.

Nota

¹ No se dispone de información suficiente para conocer si el concursante hubiera ingresado realmente o no a dicha escuela; este indicador se construyó únicamente con el

propósito de estimar si el estudiante calculó que su desempeño en el examen hubiera, al menos, sido cercano al punto de corte de las escuelas que eligió.

Referencias

- Abdulkadiroglu, A. y Sonmez, T. (2003). "School choice: A mechanism design approach", *American Economic Review*, vol. 93, núm. 3, pp. 729-47.
- Abdulkadiroglu, A.; Pathak, P.; Roth, A. y Sonmez, T. (2005). "The Boston Public School Match", *American Economic Review*, vol. 95, núm. 2, pp. 368-71.
- Allen, D. (1999). "Desire to finish college: An empirical link between motivation and persistence", *Research in Higher Education*, vol. 40, núm. 4, pp. 461-485.
- Apestequia, J. y Ballester, M. A. (2011). "Welfare of naive and sophisticated players in school choice", *Economics Working Papers 1280*, Barcelona: Department of Economics and Business, Universitat Pompeu Fabra.
- Beecher, M. y Fischer, L. (1999). "Higher school courses and scores as predictors of college success", *The Journal of College Admissions*, vol. 163, pp. 4-9.
- Bourdieu, P. (1977). "Cultural reproduction and social reproduction", en J. Karabel y A. H. Halsey (eds.), *Power and ideology in education*, Nueva York: Oxford University Press.
- Brasington, D. y Hite, D. (2012). "School choice and perceived school quality", *Economics Letters*, vol. 116, núm. 3, pp. 451-453.
- Cantalá, D. (2010). "Mercados de asignación estudiantil en la Zona Metropolitana", en Castañeda, A. (coord.), *Los grandes problemas de México*, México, DF: El Colegio de México, pp. 377-406.
- Chumacero, R.; Gómez, D. y Paredes, R. (2011). "I Would Walk 500 Miles (if it paid)", *Economics of Education Review*, vol. 30, núm. 5, pp. 1103-1114.
- Correos de México (2014). *Consulta Códigos Postales*. Disponible en <http://www.correosdemexico.gob.mx/ServiciosLinea/Paginas/ccpostales.aspx> (consultado el 15 de enero 2014).
- Cortina, J. (1993). "What is coefficient alpha: an examination of theory and applications", *Journal of Applied Psychology*, vol. 78, núm. 1, pp. 98-104.

- DeVellis R. (2003). *Scale development: theory and applications: theory and application*, Thousand Okas: Sage.
- Duncan, G., Yeung, J. W., Brooks-Gunn, J. y Smith, J. R. (1998). "How much does childhood poverty affect the life chances of children?", *American Sociological Review*, Vol. 63, pp. 406-423.
- Duncan, G.; Yeung, J. W.; Brooks-Gunn, J. y Smith, J. R. (1998). "How much does childhood poverty affect the life chances of children?", *American Sociological Review*, vol. 63, pp. 406-423.
- Fernández, N. (2004). "Hacia la convergencia de los sistemas de educación superior en América Latina", *Revista Iberoamericana de Educación*, núm. 35, pp. 39-71.
- Fernández, T.; Banegas, I. y Blanco, E. (2004). *Fundamentos teóricos de los conceptos propuestos a ser observados en los módulos permanentes del Sistema de cuestionarios de contexto para el ciclo de evaluaciones 2005-2008*, México, DF: Instituto Nacional para la Evaluación de la Educación.
- GeoNames (2014). *The GeoNames geographical database*. Disponible en: <http://download.geonames.org/export/dump/> (consultado: 17 de enero del 2014).
- Gustafsson, J. E. (1980). "Testing and obtaining fit of data to the Rasch model", *British Journal of Mathematical and Statistical Psychology*, vol. 33, pp. 205-233.
- Hair, J. F. Jr.; Anderson, R. E.; Tatham, R. L. y Black, W. C. (1998). *Multivariate data Analysis*, 5th ed., Nueva Jersey: Prentice Hall.
- Hair, J.F. Jr.; Anderson, R.E.; Tatham, R.L. y Black, W.C. (1999). *Análisis multivariante*, 5ª ed., Madrid: Prentice Hall Iberia.
- Heckman, J.; Stixrud, J. y Urzua, S. (2006). "The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior", *Journal of Labor Economics*, vol. 24, núm. 3, pp. 411-482.
- Huston, A. C. (1999). "Effects of poverty on children", en L. Balter y C. S. Tamis-LeMonda (eds.) *Child psychology: A handbook of contemporary issues*, Philadelphia: Psychology Press, pp. 391-411.
- INEGI (2010). *Censo de Población y Vivienda 2010*. Disponible en: <http://www.inegi.org.mx/lib/olap/consulta/general> (consultado: 25 de mayo de 2012).
- Lang, D.; Collins, D. y Kearns, R. (2011). "Understanding modal choice for the trip to school", *Journal of Transport Geography*, vol. 19, núm. 4, pp. 509-514.
- Masters, G. N. (1982). "A Rasch model for partial credit scoring", *Psychometrika*, vol. 47, pp. 149-174.
- McGrath, M. y Braunstein, A. (1997). "The prediction of freshmen attrition: An examination of the importance of certain demographic, academic, financial, and social factors", *College Student Journal*, vol. 31, pp. 396-408.
- Montgomery, M. (2002). "A nested logit model of the choice of a graduate business school", *Economics of Education Review*, vol. 21, núm. 5, pp. 471-480.
- Murnane, R. J.; Willett, J. B. y Levy, F. (1995). "The growing importance of cognitive skills in wage determination", *The Review of Economics and Statistics*, vol. 77, núm. 2, pp. 251-66.

- O'Shaughnessy, T. (2007). "Parental choice and school quality when peer and scale effects matter", *Economics of Education Review*, vol. 26, núm. 4, pp. 501-515.
- Pett, M. A.; Lackey, N. R. y Sullivan J. J. (2003). *Making sense of factor analysis: The use of factor analysis for instrument development in health care research*, California: Sage Publications Inc.
- Nunnally, J. y Bernstein L. (1994). *Psychometric theory*, Nueva York: McGraw-Hill Higher.
- Roth, A. (2002). "The economist as engineer: Game theory, experimental economics and computation as tools of design economics", *Econometrica*, vol. 70m núm. 4, pp. 1341-1378.
- Schulz, W. (2005). *Measuring the socio-economic background of students and its effect on achievement in PISA 2000 and PISA 2003*, documento presentado en la Reunión Anual de la American Educational Research Association, San Francisco, 7 al 11 de abril.
- SEP (2012). *Reporte de la Encuesta Nacional de Deserción en la Educación Media Superior*. México, DF: Secretaría de Educación Pública. Disponible en: http://www.sems.gob.mx/work/models/sems/Resource/10787/1/images/Anexo_6Reporte_de_la_ENDEMS.pdf (consultado: 3 de noviembre de 2013).
- Sonmez, T. (1997). "Manipulation via capacities in two-sided matching markets", *Journal of Economic Theory*, vol. 77, núm. 1, pp. 197-204.
- Streiner, D. (2003). "Starting at the beginning: an introduction to coefficient alpha and internal consistency", *Journal of Personality Assessment*. núm. 80, pp. 99-103.
- Sullivan, A. (2002). "Bourdieu and education: How useful is Bourdieu's theory for researchers?" *Netherlands' Journal of Social Sciences*, vol. 38, núm. 2, pp. 144-66.
- Thaler, R. y Sunstein, C. (2008). *Nudge: Improving decisions about health, wealth, and happiness*, Londres: Yale University Press.
- Wright, B. D. y Masters, G. N. (1982). *Rating scales analysis. Rasch measurement*, Chicago: Mesa Press.
- Zietz, J. y Joshi, P. (2005). "Academic choice behavior of high school students: Economic rationale and empirical evidence", *Economics of Education Review*, vol. 24, pp. 297-308.

Recibido: 9 de enero de 2015

Dictaminado: 13 de abril de 2015

Segunda versión: 1 de junio de 2015

Aceptado: 2 de junio de 2015

