

FACTORES RELACIONADOS CON LA EFICACIA ESCOLAR DE LAS ESCUELAS SECUNDARIAS PÚBLICAS DEL DISTRITO FEDERAL, MÉXICO

Mauricio Arce Orozco

1. OBJETIVO DEL PROYECTO DE INVESTIGACIÓN

El objetivo de la presente investigación es realizar un acercamiento para determinar los factores relacionados con la eficacia escolar de las escuelas secundarias públicas del Distrito Federal, México y determinar el nivel de eficacia de cada una de ellas y el nivel de las secundarias por zona geográfica (delegación).

Para medir la eficacia de las secundarias se utilizará los resultados del Examen Nacional de Ingreso a la Educación Media Superior (EXANI-I) desarrollado por el Centro Nacional de Evaluación para la Educación Superior (Ceneval), que es un test estandarizado referido a la norma que se emplea en el concurso de ingreso a la Educación Media Superior de México, que es aplicado a los estudiantes al egreso de la secundaria.

Este reporte parcial de resultados, muestra los primeros avances de una investigación que se realiza en el Ceneval y que es apoyada por el CONACYT con el Fondo SEP-CONACYT. Aquí se presenta un acercamiento al modelo que representa el rendimiento en habilidad verbal y matemática, empleando datos de 2004. Para el trabajo final se obtendrá el modelo para los años 2003, 2004 y 2005 (por separado) para comparar los hallazgos entre los distintos años y verificar que la eficacia escolar de las escuelas tiene que continuar a lo largo del tiempo.

Objetivos específicos

- Determinar los factores relacionados con el rendimiento escolar de habilidad verbal y matemática.
- Establecer un modelo estadístico que describa el rendimiento escolar (habilidad verbal y matemática) y el entorno socioeconómico, cultural y rendimiento previo de secundaria.
- Determinar si existe diferencia en rendimiento escolar entre secundarias y entre delegaciones donde se ubica la escuela.

2. METODOLOGÍA

Se trabajó con las bases de datos de 2004 de: a) La SEP, contienen los resultados de la evaluación de los estudiantes cuando ingresan a primer año de secundaria (instrumento de diagnóstico de alumnos de nuevo ingreso a secundaria, IDANIS); b) El CENEVAL, contienen los resultados del examen de ingreso al bachillerato (EXANI-I), junto con los datos de cuestionario contexto. Por lo tanto, se tendrá una evaluación inicial (IDANIS) y una evaluación final (EXANI-I).

Los datos para esta investigación tienen una estructura jerárquica, en donde se considera que los alumnos están agrupados en escuelas y que comparten características similares entre esos

estudiantes. A su vez, las escuelas se agrupan en delegaciones políticas dentro del Distrito Federal y que se piensa que comparten las mismas características.

Con esta información se realizó un análisis descriptivo para identificar las variables relacionadas por el rendimiento y posteriormente un análisis multinivel para establecer un modelo predictivo, estableciendo los factores que están relacionadas con el rendimiento escolar. Así como determinar si existe diferencia significativa entre los estudiantes de la misma escuela, diferencias entre las escuelas y diferencias significativas entre las distintas delegaciones en donde se ubica las secundarias.

2.1. Hipótesis

Para el desarrollo de la investigación se plantean las siguientes hipótesis:

- El entorno socioeconómico de los estudiantes afecta de forma significativa el desempeño académico de los estudiantes de secundaria.
- El nivel de habilidades y conocimientos de inicio a la secundaria es un factor que se relaciona con el rendimiento al salir de la escuela.
- La ubicación de la escuela es un factor que interviene en el rendimiento escolar.

2.2. Población objetivo

Para esta investigación se contempló estudiar a la población de egresados de las escuelas secundarias públicas del Distrito Federal, México que desean ingresar a la Educación Media Superior, por lo que sólo se incluyó a los sustentantes que estaban cursando el tercer año de secundaria en el momento del examen, excluyendo a los sustentantes egresados en años anteriores.

Los datos de estos sustentantes se cruzaron con los datos del IDANIS, eliminando los datos que no contienen información de ambos exámenes. Posteriormente, se excluyó los datos de los estudiantes que no habían registrados su datos socioeconómicos y por último, se agruparon los datos de los estudiantes por escuela, eliminando los datos de las escuelas donde existen menos de 30 estudiantes, quedando 61,495 sujetos que son objeto de este estudio de un total de 100 mil estudiantes aproximadamente.

3. ANÁLISIS DESCRIPTIVO DE LOS DATOS

Primero se realizó un análisis descriptivo para cada una de las variables que se emplearan en la investigación. De estos análisis se identificaron algunas variables que tienen relación con la medida de rendimiento, las cuales a continuación se muestra.

El 51% de los sujetos estudiados son mujeres. En la calificación global del examen, los hombres obtienen en promedio una calificación de 50.68% de aciertos, mientras que las mujeres obtienen un calificación de 47.3 % de aciertos. En habilidad verbal los hombres obtienen una calificación de 0.6 % mayor que la de las mujeres y en habilidad matemática de 6.8%.

Las variables tiempo de estudio y de tareas que los alumnos emplean fuera del horario escolar tiene una relación importante con la calificación global que obtienen en el examen. Las personas que estudian más tiempo obtienen, en promedio, una mayor calificación en el examen. Una relación similar se observa en las preguntas: horas de lectura a la semana y número de libros leídos en el último año con la calificación global del examen. La calificación menor la obtienen los estudiantes que declaran no leer, mientras que las calificaciones mayores la obtienen los estudiantes que declaran leer más de 10 horas a la semana. Esta misma relación se observa entre la cantidad de libros leídos en el

último año y la calificación global. De la misma forma, el ingreso familiar mensual es una variable que presenta relación con el nivel de rendimiento, aunque se sabe que esta variable es subrepresentada

En contraste, el número de hermanos y el índice de hacinamiento tienen una relación inversa con la calificación.

Con el objeto de evaluar las características de las casas donde viven los estudiantes, se realizaron un conjunto de 20 preguntas acerca de los bienes y servicios con los que cuenta la casa, los cuales, se agruparon en tres categorías:

1. **Servicios Básicos:** drenaje, agua entubada, alumbrado público, calles pavimentadas, recolección periódica de basura, calentador de gas para agua y cuarto propio para dormir.
2. **Bienes de entretenimiento:** automóvil familiar o propio, teléfono, teléfono celular, televisión, televisión por cable o por satélite, videograbadora y reproductor de DVD.
3. **Bienes de apoyo al estudio:** lugar exclusivo para estudiar, calculadora, computadora, diccionario o enciclopedia, suscripción a revista o periódico y conexión a Internet.

Formándose tres nuevas variables con la suma de los servicios de cada categoría que poseen en sus casas. Estas nuevas variables presentan relación positiva con la calificación global en el examen.

Otra variable que tiene relación positiva con el examen es la escolaridad de los padres. Se observó que a mayor nivel de estudios de los padres, los estudiantes obtienen mayor calificación. Por su parte, las calificaciones más altas las obtienen los estudiantes de la Secundaria Técnica, ésta obtiene una calificación global de 50.0 % de aciertos. Con una calificación un poco menor, le sigue la General con una calificación global de 48.8 %.

4. ANÁLISIS MULTINIVEL

Para identificar las variables relacionadas con la eficacia se realizará un análisis multinivel con un modelo de tres niveles:

- Primer nivel: Estudiantes
- Segundo nivel: Escuelas
- Tercer nivel: Delegaciones

En el análisis descriptivo de los datos se encontró que existe una estrecha relación entre las calificaciones y algunas de las variables. Para obtener el modelo de eficacia escolar, se utilizaran esas variables para agregarlas y determinar si son significativas en el modelo. Entre estas variables se encuentra:

- Género
- Situación laboral
- Horas de estudio a la semana
- Horas a la semana que dedica a la lectura
- Número de hermanos
- Índice de hacinamiento
- Ingreso familiar mensual
- Bienes y servicios de la casa
- Nivel de estudios de los padres

Se analizó por separado cada una de las dos habilidades y se planteó un modelo para la habilidad verbal y otro para habilidad matemática. A continuación se describe cada uno de ellos.

4.1. Modelo multinivel para Habilidad Verbal

La elaboración del modelo multinivel se inició con el modelo nulo, que es aquel donde los puntos de corte son aleatorios y no incluye ningún predictor. Este modelo se estableció como línea base para la evaluación comparada de modelos.

$$\gamma_{ij} = \beta_o + \mu_{oj} + \varepsilon_{ij}$$

Se ejecutó el modelo con el programa MLWin. Como resultado de esta primera corrida, se comprobó que el rendimiento medio de todos los estudiantes de todas las escuelas, de todas las delegaciones es de 53.43 puntos, con un error estándar de 0.56. Se trata de un valor estadísticamente significativo. La varianza entre el rendimiento medio de todas las delegaciones fue 2.95 con un error estándar de 1.7 que no era significativo, es decir, no existe variación entre las escuelas de cada delegación. Por tal razón, se procede a eliminar el tercer nivel y el nuevo modelo quedará con dos niveles, estudiantes y escuelas.

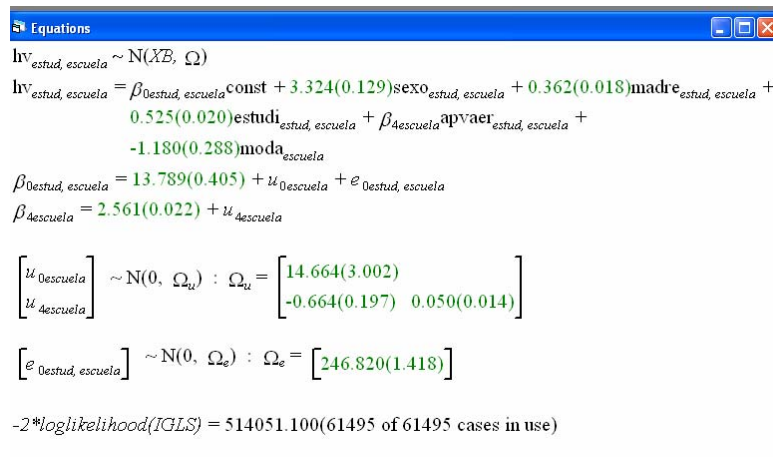
El modelo de dos niveles se corrió, teniendo el siguiente resultado: el rendimiento medio de todos los estudiantes de todas las escuelas fue de 53.45 puntos, con un error estándar de 0.31. Se trata por tanto de un valor estadísticamente significativo. La varianza entre el rendimiento medio de todas las escuelas fue de 51.02 (3.316), también estadísticamente significativo. Por último, la varianza entre los alumnos dentro de las escuelas es de 327.258 y un error estándar de 1.875 (significativo). Como consecuencia podemos decir que tiene sentido tratar de explicar la varianza observada entre escuelas y entre sujetos dentro de las escuelas.

El siguiente paso consistió en añadir diferentes predictores al modelo. El primero que se agregó fue la variable sexo para el primer nivel. Se ejecutó el modelo y se obtienen los resultados. En este modelo se observó que la varianza de la variable sexo era de 0.659 con un error estándar de 0.66, la cual resultó no significativa, del mismo modo que la covarianza entre sexo y el intercepto, por lo que se fijó estos parámetros en cero y ejecutamos nuevamente el modelo. Teniendo que los parámetros fueron significativos, además de que el valor del intercepto disminuyó, pasando de 53.45 a 52.11 puntos. Esto se debe a que ahora ese parámetro ya no es la medida de todos los alumnos de todas las escuelas, sino la media de todos los alumnos para los que la variable sexo es igual a cero, es decir, las mujeres.

El parámetro asociado al sexo toma un valor de 2.89, que es la diferencia entre el rendimiento medio de los hombres y las mujeres. El rendimiento en habilidad verbal base es de 52.11, si el estudiante es hombre se le suma 2.89 puntos. También se observa que la mayor variación del rendimiento esta entre los estudiantes de cada escuela, con un valor de 325.2, mientras que la varianza entre las escuelas es de 51.27 puntos.

Después se añadió la variable años de estudio de la madre. Ejecutando el modelo se observó que el parámetro de la covarianza de esta variable y el intercepto no fue significativa, por lo que se fija en cero y se ejecuta nuevamente el modelo. Después, se observó que todos los parámetros son significativos y el valor de la verosimilitud disminuyó. También se observó que el valor del intercepto disminuyó a 46.7 puntos. Este valor ahora representa el rendimiento promedio de los estudiantes mujeres cuya madre no tiene ningún año de estudio.

Posteriormente se añade la variable “estudi” que representa las horas que el estudiante estudia a la semana. La varianza y la covarianza resultaron no significativas, por lo que se fijaron a cero. Posteriormente, se añadieron las variables apvaer (habilidad verbal del IDANIS) y “moda” modalidad de la secundaria, con los siguientes resultados.



```

Equations
hv_{estud, escuela} ~ N(\lambda B, \Omega)
hv_{estud, escuela} = \beta_{0estud, escuela} \text{const} + 3.324(0.129)\text{sexo}_{estud, escuela} + 0.362(0.018)\text{madre}_{estud, escuela} +
0.525(0.020)\text{estudi}_{estud, escuela} + \beta_{4escuela}\text{apvaer}_{estud, escuela} +
-1.180(0.288)\text{moda}_{escuela}
\beta_{0estud, escuela} = 13.789(0.405) + u_{0escuela} + e_{0estud, escuela}
\beta_{4escuela} = 2.561(0.022) + u_{4escuela}
[u_{0escuela}, u_{4escuela}] ~ N(0, \Omega_u) : \Omega_u = [14.664(3.002), -0.664(0.197); -0.664(0.197), 0.050(0.014)]
[e_{0estud, escuela}] ~ N(0, \Omega_e) : \Omega_e = [246.820(1.418)]
-2*loglikelihood(IGLS) = 514051.100(61495 of 61495 cases in use)

```

4.2. Modelo de dos niveles final de HV

Se trató de añadir más variables al modelo, resultando algunas no significativas. Por otro lado, en el caso de algunas variables resultó que el nuevo modelo no era mejor que el anterior, por lo que el modelo final es el que se presenta en la figura anterior. Se observa que el valor del intercepto es de 13.8 puntos que representa el rendimiento promedio de los estudiantes mujeres cuyas madres tienen cero años de estudio, que no estudian ninguna hora a la semana, que tienen una calificación de 0 puntos en habilidad verbal cuando ingresaron a la secundaria y que provienen de escuela secundaria técnica. A esta calificación se le suman puntos de acuerdo a lo siguiente:

- Si el estudiante es hombre se le suman 3.32 puntos
- Se suman 0.36 puntos por cada año de estudio de la madre
- 0.525 puntos por cada hora de estudio a la semana
- 2.6 puntos por cada punto en la calificación de habilidad verbal al ingreso de la secundaria y
- Se le resta -1.18 puntos si el estudiante proviene de secundaria general.

4.3. MODELO MULTINIVEL PARA HABILIDAD MATEMÁTICA

Por otra parte, se analizó el rendimiento en habilidad matemática. Se siguió el mismo proceso que para habilidad verbal. Partiendo del modelo nulo y agregando diferentes variables, llegando al siguiente resultado:

El modelo representa el rendimiento en habilidad matemática, el intercepto (β_0) representa la media de las estudiantes: mujeres, con madres que no tienen ningún año de escolaridad, que no estudian ninguna hora a la semana, que no tienen ningún servicios de apoyo al estudio, que obtuvieron una calificación de cero en el habilidad matemáticas del examen de ingreso a la secundaria (IDANIS) y que provienen de escuelas técnicas

MODELO DE DOS NIVELES FINAL DE HABILIDAD MATEMÁTICA

Equations

$$hm_{estud, escuela} \sim N(\beta, \Omega)$$

$$hm_{estud, escuela} = \beta_{0escuela}const + \beta_{1escuela}sexo_{estud, escuela} + 0.321(0.020)madre_{estud, escuela} + \beta_{3escuela}estudi_{estud, escuela} + 0.707(0.053)ser_{estud, escuela} + \beta_{5escuela}apmat_{estud, escuela} - 1.141(0.423)moda_{escuela}$$

$$\beta_{0escuela} = 19.437(0.475) + u_{0escuela} + e_{0estud, escuela}$$

$$\beta_{1escuela} = 6.392(0.156) + u_{1escuela}$$

$$\beta_{3escuela} = 0.687(0.024) + u_{3escuela}$$

$$\beta_{5escuela} = 2.381(0.032) + u_{5escuela}$$

$$\begin{bmatrix} u_{0escuela} \\ u_{1escuela} \\ u_{3escuela} \\ u_{5escuela} \end{bmatrix} \sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} 18.233(3.028) & & & \\ 2.605(1.042) & 1.648(0.707) & & \\ 0 & 0 & 0.028(0.013) & \\ -1.015(0.269) & -0.239(0.104) & 0 & 0.165(0.029) \end{bmatrix}$$

$$[e_{0estud, escuela}] \sim N(0, \Omega_e) : \Omega_e = [300.571(1.738)]$$

-2*loglikelihood(TGLS) = 526665.800(61495 of 61495 cases in use)

Al puntaje del intercepto 19.43 se le suman puntos de acuerdo a la característica del estudiante.

- Si es hombre se le suman 6.4 puntos
- Por cada año de estudios de la madre se le suman 0.32 puntos
- Por cada hora de estudio a la semana se le suman 0.69
- Por cada servicio de apoyo al estudio que se tenga se sumaran 0.71 puntos
- Por cada punto en la calificación de matemáticas en el IDANIS se le suman 2.38 puntos
- Si proviene de secundaria general se le restan 1.14 puntos.

5. CONCLUSIONES

En este estudio se ha aportado una fuerte evidencia empírica de que existe una importante influencia de las variables género del sustentante, escolaridad de la madre, horas de estudio y nivel de rendimiento previo del estudiante en el rendimiento académico de los estudiantes en habilidad verbal y habilidad matemática.

Las hipótesis planteadas en este estudio fueron:

- El entorno socioeconómico de los estudiantes afecta de forma significativa el desempeño académico de los estudiantes de secundaria.
- El nivel de habilidades y conocimientos de inicio a la secundaria es un factor que se relaciona con el rendimiento al salir de la escuela.
- La ubicación de la escuela es un factor que interviene en el rendimiento escolar.

En esta investigación se observaron suficientes evidencias empíricas para confirmar la primera y la segunda hipótesis, ya que se observó algunas que variables socioeconómicas están relacionadas con el rendimiento escolar, como se pudo demostrar en el modelo multinivel resultante. Sin embargo, la tercera hipótesis no se pudo confirmar.

Para continuar con la investigación, se agregará información a nivel de aula, escuela y delegación que se añada al modelo. Así mismo, se realizará estos análisis para varios años y de esta forma verificar que las escuelas eficaces conserven su nivel a través del tiempo.