



**Aplicación de modelos de ecuaciones estructurales en la
medición del nivel de satisfacción estudiantil: comparación de
tres métodos de estimación**

Ramón Álvarez - Elena Vernazza
Noviembre 2014

Documentos de Trabajo

Serie DT (14 / 02) - ISSN : 1688-6453

Aplicación de modelos de ecuaciones estructurales en la medición del nivel de satisfacción estudiantil: comparación de tres métodos de estimación

Ramón Álvarez ¹

Instituto de Estadística - Facultad de Ciencias Económicas y de Administración - Udelar.

Elena Vernazza ²

Instituto de Estadística - Facultad de Ciencias Económicas y de Administración - Udelar.

RESUMEN

En este trabajo se estudian las propiedades psicométricas de un instrumento propuesto para medir la satisfacción estudiantil para los cursos superiores de la Universidad de Beira Interior (Portugal), para luego ver los resultados que surgen de aplicarlo para el caso de la Facultad de Ciencias Económicas y de Administración, Udelar (Uruguay).

El indicador propuesto para medir el nivel de satisfacción estudiantil considera relaciones de causa-efecto entre algunas variables que son consideradas como “antecedentes” y otras como “consecuencia” de la *satisfacción*, dentro de las que están *expectativas* de los alumnos, la *imagen* que tienen de la facultad, la *calidad* de la enseñanza y servicios, y el *valor percibido*.

Los datos utilizados para la aplicación presentada en este trabajo provienen de una encuesta aplicada sobre una muestra probabilística de estudiantes de la facultad, en el año 2009.

El cuestionario aplicado, presenta 9 bloques de preguntas, uno con características sociodemográficas de los estudiantes y otros con variables que forman parte del modelo ECSI (European Customer Satisfaction Index) y serán las utilizadas como insumos para el cálculo del índice de satisfacción estudiantil.

Los resultados, presentados para un modelo con 22 variables observables y 7 constructos no observables, se comparan para tres métodos de estimación: máxima verosimilitud (MV), mínimos cuadrados generalizados (MCG) y mínimos cuadrados parciales (MCP - PLS).

Palabras claves: Análisis factorial, índice de satisfacción, modelos de ecuaciones estructurales, mínimos cuadrados parciales.

¹ramon@iesta.edu.uy

²evernazza@iesta.edu.uy

1. Introducción

Conocer el nivel de satisfacción de los clientes, con determinado servicio al cual acceden, resulta fundamental ya que en función de esto se podrán tomar decisiones que tengan como objetivo primordial mantener ó mejorar, en caso de que sea necesario, aquellos aspectos que se entiende determinan la “satisfacción”.

En un intento de vincular esta idea con la educación universitaria, se toma lo propuesto por Alves y Raposo (Alves and Raposo, 2005), quienes plantean: “Sólo con la satisfacción de los alumnos se podrá alcanzar el éxito escolar, la permanencia de los estudiantes en la institución y, sobre todo, la formación de una valoración positiva boca a boca. En este sentido, es extremadamente importante encontrar formas fiables de medir la satisfacción del alumno en la enseñanza universitaria, permitiendo así a las instituciones de enseñanza conocer su realidad, compararla con la de los otros competidores y analizarla a lo largo del tiempo”.

En este trabajo se considera a los estudiantes universitarios que concurren a la Facultad de Ciencias Económicas y Administración, FCCEEyA - UdelAR, como “clientes” y se entiende que el “servicio” que se les brinda es el de la educación a nivel terciario. Sin embargo, tal como establecen Blanco y Blanco (Blanco and Blanco Peck, 2007) ³ no se pueden dejar de lado los valores y metas de la Universidad como institución, es decir, no debe perderse la visión humana de los estudiantes que forman parte de ella. Logrando esto, se evitará considerar al modelo de educación como un modelo industrial, donde se estaría considerando a los estudiantes como simples productos del sistema.

Lograr conocer la dimensión de la satisfacción de los estudiantes con la facultad a la cual concurren, permitirá identificar aspectos tanto positivos como negativos, siendo estos últimos fundamentales a la hora de determinar estrategias de mejora de la educación.

La información necesaria para poder evaluar y entender por un lado, qué conceptos se asocian a la satisfacción y por otro, cómo se establecen la interrelaciones entre estos conceptos, se obtiene a través de la aplicación de un cuestionario formado por una serie de bloques de preguntas que conforman el modelo ECSI (European Customer Satisfaction Index); sobre este instrumento con la ayuda del Análisis Factorial y, más precisamente, de los Modelos de Ecuaciones Estructurales que se presentan en las secciones 2.1 y 2.2, se logra poner de manifiesto los componentes de la satisfacción.

El presente trabajo se estructura en 6 secciones. En primera instancia se presenta la

³citando a Gaitán y López, 1999. “La medición de la calidad de servicios en la educación universitaria”

metodología utilizada, se hace referencia a los principales aspectos del análisis factorial y se pone especial énfasis en la presentación de los modelos de ecuaciones estructurales. Las secciones 3 y 4 muestran la aplicación y los principales resultados obtenidos. A continuación, en la sección 5 se plantean las comparaciones de los resultados obtenidos por tres métodos de estimación (MV - MCG -MCP) y, por último se presentan las principales conclusiones y pasos a seguir.

2. Metodología

2.1. Análisis factorial (AF)

El análisis factorial tiene su origen a principios del siglo XX, cuando Karl Pearson y Charles Spearman muestran su interés en comprender las distintas dimensiones que conforman la inteligencia humana. Desde ese momento, y hasta la actualidad, se han realizado aplicaciones en diversas áreas de estudio, entre las que se destacan fundamentalmente los trabajos realizados en el área de la psicometría. Esta disciplina se encarga de asignar un valor numérico a las diversas características psicológicas de las personas, engloba tanto la teoría como la elaboración de pruebas y tests. Incluye, además, la formulación y aplicación de procedimientos estadísticos que permiten determinar si una prueba efectivamente resulta válida para la medición de una variable o conducta psicológica previamente definida.

El principal objetivo del análisis factorial (Dillon y Goldstein, 1984),(Peña, 2002),(Revelle, 2004) es reducir las dimensiones del problema en estudio, descartando información redundante. A partir de una gran cantidad p de variables observadas x_1, \dots, x_p se construye una menor cantidad m de variables no observadas (latentes), denominadas factores f_1, \dots, f_m . Estas últimas resultan de combinar linealmente las variables observadas. Tal como establece Blanco (Blanco, 2006), además de buscar eliminar la información redundante, se debe lograr que esa simplificación implique la menor pérdida posible de información, es decir, que la nube original, al ser proyectada sobre un espacio de menor dimensión, se deforme lo menos posible. Esto implicará que la distancia entre los puntos originales y su proyección sea mínima.

Formulación del AF

La representación matricial del modelo factorial general es la siguiente:

$$x = \mu + \Lambda f + \varepsilon \quad (1)$$

donde:

- μ es el vector de medias de x

- $x_{p \times 1}$ vector que contiene las p variables observadas
- Λ_p matriz de coeficientes que describen cómo los factores f afectan a las variables observadas x . Matriz de cargas.
- $f_{m \times 1}$ vector de variables latentes/factores, tal que: $f \sim (0, I)$
- $\varepsilon_{p \times 1}$ vector de perturbaciones (no observadas), tal que: $\varepsilon \sim (0, \Psi)$. Los errores se suponen independientes por lo que Ψ resulta una matriz diagonal
- $cor(f, \varepsilon) = 0$

Si además se supone que tanto los factores como las perturbaciones siguen una distribución normal multivariada, $f \sim N_m(0, I)$ y $\varepsilon \sim N_p(0, \Psi)$, se deduce que el vector de variables observadas x también tiene distribución normal (por ser combinación lineal de variables aleatorias normales) y que su vector de medias es μ , por lo tanto se tiene: $x \sim N_p(\mu, \Sigma)$.

A partir de la ecuación (1) se tiene que, dada una muestra aleatoria de tamaño n , cada observación x_{ij} de la muestra queda determinado por:

$$x_{ij} = \mu_j + \lambda_{j1}f_{1i} + \dots + \lambda_{jm}f_{mi} + \varepsilon_{ij} \quad (2)$$

con: $i = 1, \dots, n$ y $j = 1, \dots, p$.

Los primeros 2 a $m + 1$ términos, reflejan el efecto de los m factores, y el último muestra una perturbación específica de cada observación.

De forma de simplificar la notación, de aquí en más se entenderá que las variables están centradas, es decir que cada una de ellas tiene media 0 y que, por lo tanto, el vector de medias del vector x es el vector nulo.

En función de esto si se consideran todas las ecuaciones, la matriz de datos X_n puede plantearse como:

$$X_{n \times p} = F_{n \times m} \Lambda'_{m \times p} + \varepsilon_{n \times p} \quad (3)$$

A partir de los supuestos realizados, se tiene: $cov(f, x) = \Lambda$, ya que:

$$cov(f, x) = E(xf') - \underbrace{E(x)}_{=0} \underbrace{E(f')}_{=0} = E((\Lambda f + \varepsilon)f') = E(\Lambda f f') + \underbrace{E(\varepsilon f')}_{=0} = \Lambda \underbrace{E(f f')}_{=I}$$

De forma equivalente, para cada i, j , se tiene: $\lambda_{i,j} = cov(x_i, f_j)$ y si se trabaja con las variables estandarizadas, la igualdad es: $\lambda_{i,j} = cor(x_i, f_j)$.

Por otra parte, la matriz de varianzas y covarianzas de las variables observadas (Σ) puede ser vista como la suma de otras dos matrices, verificando:

$$\Sigma = \Lambda\Lambda' + \Psi \quad (4)$$

donde: $\Lambda\Lambda'$ es una matriz simétrica, de rango m que contiene la parte de varianza común al conjunto de las x , y Ψ es una matriz diagonal que contiene la parte de varianza específica de cada variable. Por lo tanto, la varianza de cada variable observada x_i , será:

$$\sigma_i^2 = \sum_{j=1}^m \lambda_{ij}^2 + \psi_i \quad (5)$$

El primer término, que refiere a la proporción de variabilidad de x_i , explicada por los factores comunes, se denomina *comunalidad* y se denota h_i^2 . El segundo término refiere a la proporción de variabilidad específica de cada x_i , y se denomina *unicidad*. Siguiendo esta notación, se tiene: $\sigma_i^2 = h_i^2 + \psi_i$.

El análisis factorial puede ser aplicado como una herramienta exploratoria (análisis factorial exploratorio - AFE) o como un modelo para contrastar teorías (análisis factorial confirmatorio - AFC). El AFE, intenta reducir el número inicial de variables, en un conjunto de menor dimensión de factores. Éstos se determinan en función de algún criterio estadístico, que por lo general tiende a optar por la solución factorial que logra explicar la mayor cantidad de varianza original. Por otra parte, en el caso del AFC el número de factores se supone conocido a priori y se establecen restricciones sobre los elementos de la matriz de cargas. Por ejemplo, algunos pesos pueden ser fijados en cero o se puede determinar que haya dos pesos iguales (Peña, 2002),(Revelle, 2004).

Como un caso particular dentro del análisis factorial confirmatorio se encuentran los *modelos de ecuaciones estructurales* que se presentan a continuación.

2.2. Modelos de ecuaciones estructurales

Este tipo de modelos pueden ser vistos, fundamentalmente, de dos maneras. Por un lado, pueden ser enmarcados en el ámbito de los modelos de regresión, con ciertas particularidades que los diferencian de los modelos de regresión clásicos y, por otro, pueden ser

vistos como una técnica de análisis factorial que permite establecer relaciones entre los factores.

De modo simplificado, podría entenderse que en los modelos de ecuaciones estructurales se presentan relaciones *causales* entre, por un lado, un conjunto de variables observables y por otro, variables tanto observables como no observables.

A partir de esto, y recordando además que estos modelos se presentan en el contexto del análisis factorial confirmatorio, es que resulta fundamental establecer de forma clara el concepto de *causalidad*, ya que justamente es esta relación la que este tipo de modelos intentan *confirmar*.

Tomando como referencia lo propuesto por Casas Guillen (Casas Guillén, 2008)⁴ podríamos decir que existe una relación de causalidad entre la variable X y la variable Y y, más precisamente, que X causa a Y si cada vez que sucede X , sucede Y , y nunca se da Y sin que previamente se haya dado X .

Los modelos de ecuaciones estructurales presentan la particularidad de que una variable puede ser causada por otra variable del sistema y a la vez, dentro del mismo modelo, ser causa de otra variable.

Como se explicitó previamente, existen fundamentalmente dos tipos de relaciones presentadas en los modelos de ecuaciones estructurales. Por un lado, se establecen relaciones entre variables no observadas y, por otro, relaciones entre estas variables, y variables observadas, lo que da lugar a dos submodelos: *modelo estructural* y *modelo de medida*, respectivamente

Si bien la presentación de éstos, se plantea en términos de ecuaciones, por lo general, se agrega una representación gráfica que permite visualizar mejor las relaciones entre variables.

Previo a presentar los pasos a seguir al trabajar con este tipo de modelos, se establecen las características más relevantes de las variables que los conforman. Tanto las variables observadas como las no observadas, pueden ser de naturaleza endógena o exógena. En lo que refiere a variables latentes, de aquí en adelante se entenderá por variable endógena, a aquella variable cuyas *causas* están presentes en el modelo, éstas estarán siempre acompañadas de un término de error/perturbación. Por otra parte, una variable será tratada como exógena cuando no reciba efecto alguno de ninguna de las variables que forman parte del modelo. Éstas serán manejadas como “libres de error”. En cuanto a las variables observadas, se dirá que una variables es exógena, cuando las variables que la causan sean

⁴Citando a Bisquerra, R. (1989), en “Introducción conceptual al análisis multivariable. Vol. II, PPU, Barcelona.”

variable latentes exógenas, mientras que se estará frente a una variable endógena cuando las causas de ésta sean de naturaleza también endógena. Por último se presentan los errores, variables aleatorias no observables que recojen aquellos efectos sobre las variables dependientes, que el modelo no logra captar.

Una vez establecidas las principales características que pueden presentar las variables que se manejarán de aquí en adelante, y las posibles relaciones que podrán existir entre ellas, se presentan los detalles de los MES.

A modo de resumen, se puede decir que las principales etapas que conforman el proceso de determinar/confirmar las relaciones existentes entre determinadas variables, a partir de la implementación de *modelos de ecuaciones estructurales*, son las que se presentan en la figura 1.

2.2.1. Especificación: Modelo estructural

El modelo estructural es el submodelo, dentro de los MES, que captura las relaciones existentes entre las variables no observables, también denominadas variables latentes, constructos o factores.⁵

En formato matricial, podemos representar estos modelos de la siguiente manera:

$$\beta\eta = \Gamma\xi + \zeta \quad \Rightarrow \quad \eta = B\eta + \Gamma\xi + \zeta \quad (6)$$

donde, considerando un modelo con m variables latentes endógenas y k variables latentes exógenas, se tiene:

- β (beta) matriz, de dimensión m , de pesos β que determinan la relación entre dos variables latentes endógenas
- η (eta) vector, de dimensión $m \times 1$, de variables latentes endógenas
- Γ (gamma) matriz, de dimensión m , de pesos γ que determinan la relación entre una variable endógena y una exógena, ambas latentes
- ξ (xi) vector, de dimensión $k \times 1$, de variables latentes exógenas
- ζ (zeta) vector, de dimensión $m \times 1$, de términos de error/perturbación.

⁵De aquí en adelante, utilizaremos cualquiera de estos términos indistintamente

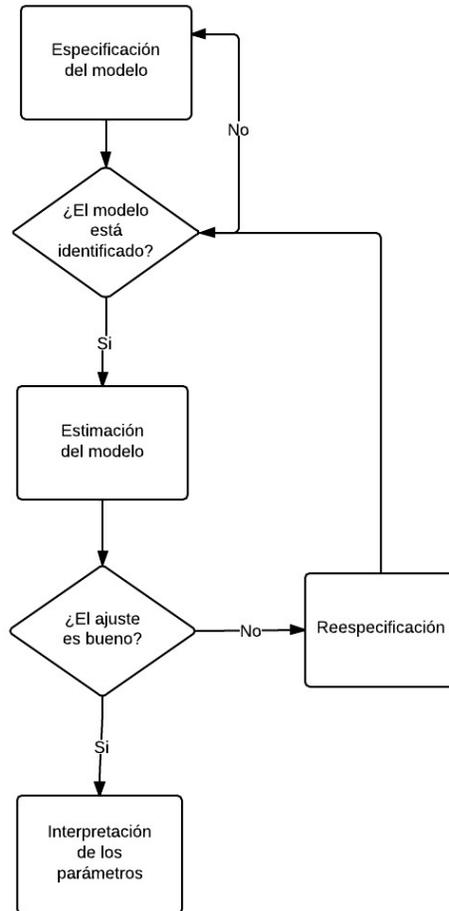


Figura 1: Pasos a seguir al trabajar con MES

También forman parte de este modelo, las matrices Φ (phi) y Ψ (psi), que representan la matriz de correlaciones entre las variables latentes exógenas (ξ) y la matriz de correlaciones entre los errores de las variables latentes endógenas (ζ), respectivamente.

A modo de ejemplo, se consideran tres variables latentes endógenas η_1, η_2, η_3 y una variable latente exógena ξ_1 y lo que se desea es confirmar que: ξ_1 y η_2 *causan* a η_1 , ξ_1 y η_3 *causan* a η_2 , y ξ_1 y η_1 *causan* a η_3 . La representación analítica del modelo es la siguiente:

$$\begin{cases} \eta_1 = \gamma_1 \xi_1 + \beta_1 \eta_2 + \zeta_1 \\ \eta_2 = \gamma_2 \xi_1 + \beta_2 \eta_3 + \zeta_2 \\ \eta_3 = \gamma_3 \xi_1 + \beta_3 \eta_1 + \zeta_3 \end{cases} \quad (7)$$

2.2.2. Especificación: Modelo de medida

En el modelo de medida se establecen las relaciones que existen entre los factores y las variables observables. Se presentan, por separado, las relaciones entre las variables exógenas y las endógenas, lo que determina dos submodelos. La expresión matricial para el modelo de medida, para las variables exógenas, queda determinado por:

$$X = \Lambda_x \xi + \delta \quad (8)$$

donde, considerando un modelo con k variables latentes y q variables observables, se tiene:

- X vector, de dimensión $qx1$, de variables observables
- Λ_x (lambda) matriz, de dimensión qxk , de pesos λ que determinan la relación entre cada x y cada ξ
- ξ (xi) vector, de dimensión $kx1$, de variables latentes exógenas
- δ (delta) vector, de dimensión $qx1$, de términos de error/perturbación

La matriz Θ_δ también forma parte de este submodelo. Ésta es la matriz de covarianzas entre los errores de las variables exógenas observadas (δ). Los errores δ se suponen incorrelacionados, por lo que la matriz Θ_δ resulta una matriz diagonal.

Para las variables endógenas, la expresión matricial para el modelo de medida, es la siguiente:

$$Y = \Lambda_y \eta + \epsilon \quad (9)$$

donde, considerando un modelo con m variables latentes y p variables observables, se tiene:

- Y vector, de dimensión $px1$, de variables observables
- Λ_y (lambda) matriz, de dimensión pxm , de pesos λ que determinan la relación entre cada y y cada η
- η (eta) vector, de dimensión $mx1$, de variables latentes endógenas
- ϵ (epsilon) vector, de dimensión $px1$, de términos de error/perturbación

La matriz Θ_ϵ también forma parte de este submodelo. Ésta es la matriz de covarianzas entre los errores de las variables endógenas observadas (ϵ). Los errores ϵ se suponen incorrelacionados, por lo que la matriz Θ_ϵ resulta una matriz diagonal.

Continuando con el ejemplo presentado para el modelo estructural, se agregan dos variables exógenas x_1, x_2 y seis endógenas $y_1 \dots y_6$, todas observables. Se supone que x_1 y x_2 son causadas por ξ_1 , y_1 y y_2 por η_1 , y_3 y y_4 por η_2 y y_5 y y_6 por η_3 . Las ecuaciones que determinan este modelo son:

$$\begin{cases} x_1 = \lambda_1 \xi_1 + \delta_1 \\ x_2 = \lambda_2 \xi_1 + \delta_2 \end{cases} \quad (10)$$

$$\begin{cases} y_1 = \lambda_3 \eta_1 + \epsilon_1 \\ y_2 = \lambda_4 \eta_1 + \epsilon_2 \\ y_3 = \lambda_5 \eta_2 + \epsilon_3 \\ y_4 = \lambda_6 \eta_2 + \epsilon_4 \\ y_5 = \lambda_7 \eta_3 + \epsilon_5 \\ y_6 = \lambda_8 \eta_3 + \epsilon_6 \end{cases} \quad (11)$$

2.2.3. Representación gráfica de los MES

Tal como se mencionó previamente, por lo general la especificación analítica de los modelos de ecuaciones estructurales se acompaña de una representación gráfica. Lo más frecuente es que ésta sea a través de los diagramas de senderos o ruta (“path diagrams”)(Kline, 2011). Este tipo de diagramas se basan en algunas convenciones que resulta conveniente explicitar, de forma tal que su “lectura” resulte sencilla.

- Variables observables: \square
- Variables latentes: \circ
- Efecto, directo, de una variable sobre otra: \longrightarrow . Al inicio de la flecha se encuentra la variable independiente y al final, la dependiente. Los errores también se vinculan con las variables dependientes a través de estas flechas.
- Covarianza/correlación entre variables: \longleftrightarrow

Siguiendo estas convenciones, el modelo estructural presentado a modo de ejemplo, en la subsección 2.2.1 con las tres variables latentes queda representado gráficamente en la figura 5.

Por su parte, el modelo de medida presentado, a modo de ejemplo, en la subsección 2.2.2 queda especificado gráficamente en el diagrama que se muestra en la figura 3 y, finalmente, el modelo MES que combina los submodelos de medida y estructural ejemplificados tiene el diagrama de senderos que se presenta en la figura 4.

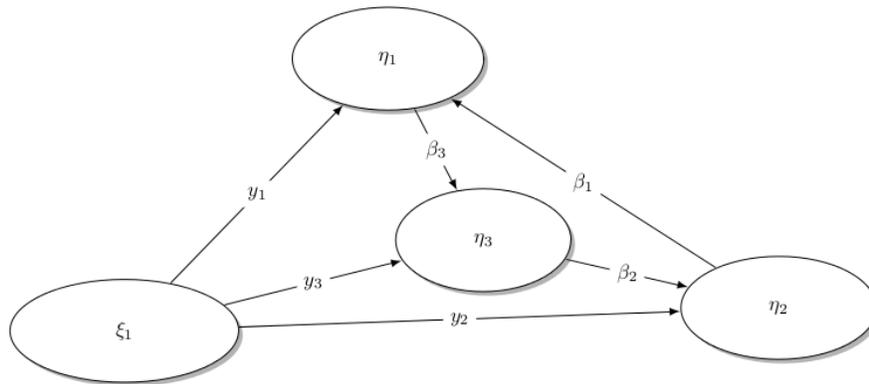


Figura 2: Modelo estructural

Una vez que quedan explicitadas por un lado, las relaciones que se desean confirmar a nivel de variables latentes (modelo estructural) y por otro, aquellas relaciones que se pretende determinar entre las variables observables y las latentes (modelo de medida), culmina el primer paso en el proceso de realizar un estudio de relaciones entre variables, a través de la utilización de *modelos de ecuaciones estructurales*, la “especificación” del modelo.

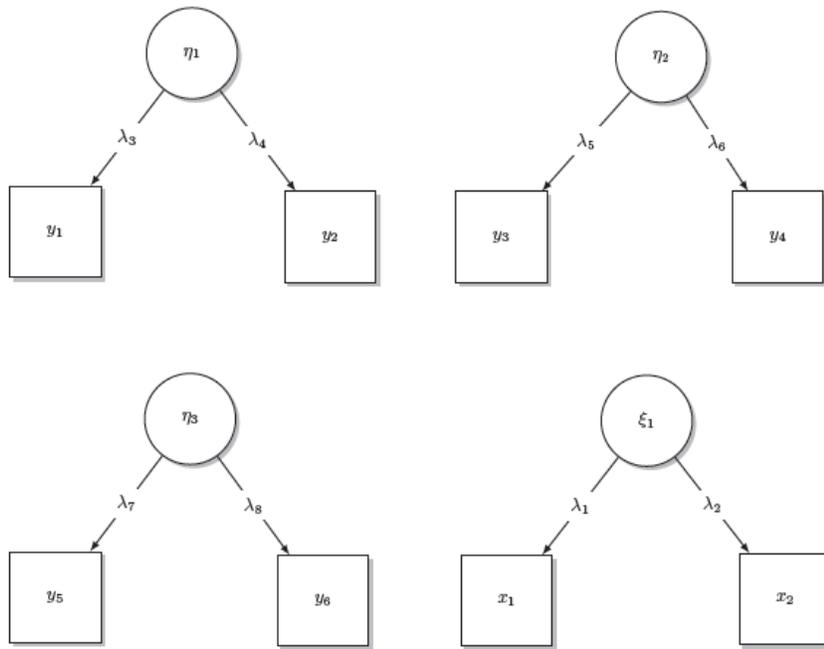


Figura 3: Modelo de medida

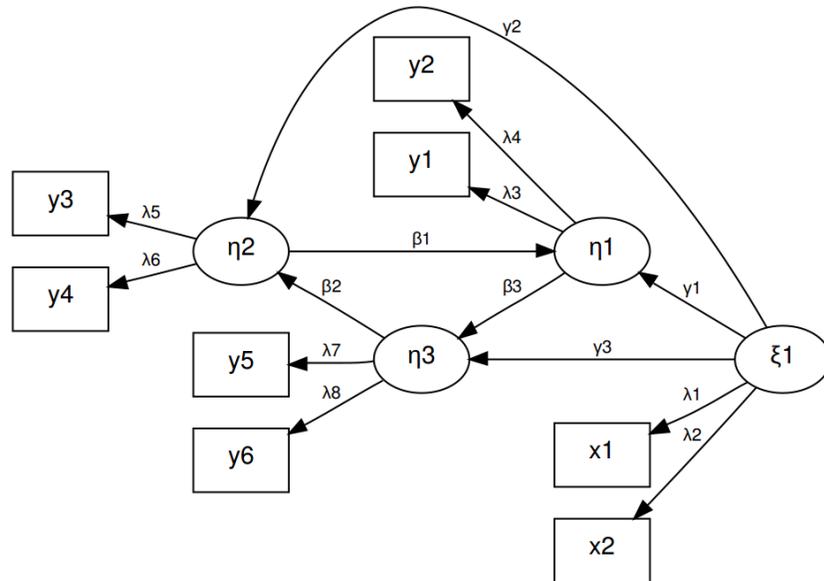


Figura 4: MES para ejemplo presentado

2.2.4. Estimación

Una vez que finaliza el proceso de identificación del modelo se pasa a la etapa que consiste en recolectar/depurar los datos y prepararlos para realizar la “estimación” del modelo, lo cual implica obtener una estimación puntual para cada uno de los parámetros involucrados en los submodelos de medida y estructural. Existen varios métodos de estimación, entre los que se destacan:

- Máxima verosimilitud (MV)
- Mínimos cuadrados generalizados (MCG)
- Mínimos cuadrados parciales (MCP)

En este trabajo se hará especial énfasis en el método de MCP.

2.2.5. Mínimos cuadrados parciales

El método a presentar es el de mínimos cuadrados parciales (PLS - MCP). La principal ventaja de este método es que, al igual que los MCG, éste no presenta restricciones distribucionales.

Retomando la notación típica de los modelos de ecuaciones estructurales, y tratando de vincularla con la terminología propia de los PLS, se establecen las siguientes ecuaciones, con sus correspondientes parámetros a estimar:

- Modelo estructural (Inner Model): Los parámetros a estimar son los $\beta's$ y $\gamma's$, también denominados “path coefficients”, de la ecuación $\beta\eta = \Gamma\xi + \zeta$
- Modelo de medida (Outer Model): Los parámetros a estimar son los denominados “loadings” λ de las ecuaciones:
 - $X = \Lambda_x\xi + \delta$
 - $Y = \Lambda_y\eta + \epsilon$

Otro elemento importante que se incorpora al considerar este método de estimación, es el hecho de que las variables latentes, justamente por su condición de no observables, deben ser presentadas/especificadas de forma tal que resulte posible introducirlas como variables a explicar o explicadas tanto en el modelo de medida como en el modelo estructural. Es para esto que se introduce la idea de “score” y “outer weights” que determinan que las variables latentes son estimadas a través de una combinación lineal de las variables observadas con las que se vincula, tal como se presenta en las siguientes ecuaciones:

$$\hat{\eta}_j = \sum_k w_{jk} Y_{jk} \quad (12)$$

$$\hat{\xi}_l = \sum_m w_{lm} X_{lm} \quad (13)$$

donde las denominaciones “score” y “outer weight” hacen referencia a $\hat{\eta}_j$ y $\hat{\xi}_l$, y w respectivamente.

Siguiendo lo propuesto por Sanchez (Sanchez, 2012), una vez definidos todos estos elementos, se puede decir que el procedimiento de estimar un modelo por mínimos cuadrados parciales consiste fundamentalmente en:

- Obtener los pesos que permitan determinar $\hat{\eta}_j$ y $\hat{\xi}_l$
- Estimar todos los elementos de β y Γ
- Obtener los λ de los dos submodelos que determinan el modelo de medida

El primer paso consiste en determinar los w , lo cual se logra implementando el siguiente proceso iterativo:

1. Se determinan valores arbitrarios para w , por lo general se inicia con $w = 1$
2. Se realiza una aproximación “externa” para las variables latentes
3. Se obtienen nuevos pesos
4. Se realiza una aproximación “interna” para las variables latentes
5. Se calculan nuevos valores para w

Se repiten los pasos 2 a 5, hasta lograr convergencia.

A continuación se presenta en detalle cada uno de estos pasos. De manera de simplificar la notación, de acá en adelante, en esta subsección no se diferenciarán las variables endógenas y exógenas. Las variables latentes se denotarán como VL y las observadas como VO , sin importar su naturaleza.

El primer paso establece que, partiendo de $\tilde{w} = 1$, se obtendrá una estimación de las variables latentes a partir de combinar linealmente las variables observadas con las que se relaciona, lo cual puede ser representado en formato matricial como:

$$\widehat{VL}_k = VO_k \tilde{w}_k$$

En el paso siguiente se consideran las relaciones entre las variables latentes (por ejemplo VL_i y VL_j) y en función de éstas se generan nuevos pesos, por ejemplo, e_{ij} . Esta nueva estimación quedará determinada de la siguiente manera:

$$\widehat{VL}_j^* = \sum_i e_{ij} \widehat{VL}_i \quad (14)$$

donde VL_i son todas las variables latentes con las que se relaciona la variable latente VL_j . Tal como establece Sanchez (Sanchez, 2012) existen tres posibles esquemas para determinar los pesos e_{ij} :

- **Centroide:** Considera únicamente el sentido de la correlación entre las variables latentes, es decir: $e_{ij} = \text{signo}(\text{cor}(\widehat{VL}_i, \widehat{VL}_j))$ si VL_i y VL_j se comunican, y 0 en otro caso.
- **Factorial:** Considera, además del signo, la magnitud de la correlación. En función de esto, los pesos e_{ij} se definen como: $e_{ij} = \text{cor}(\widehat{VL}_i, \widehat{VL}_j)$ si VL_i y VL_j se comunican, y 0 en otro caso.
- **De senderos:** Considerar la relación que existe entre las variables latentes, diferenciando a las variables en dependientes e independientes, según corresponda. Si la variable VL_i causa a la variable VL_j , entonces los pesos e_{ij} se corresponderán con los coeficientes de regresión que se obtienen al regresar VL_j sobre las VL_i que la causan. Si la relación de causalidad es la opuesta, es decir, VL_j causa a VL_i , este esquema coincide con el esquema **factorial**.

Una vez que se obtienen los pesos e_{ij} , se procede a obtener la estimación “interna” de las variables latentes, que consiste en sustituir los pesos obtenidos, en el paso anterior, en la ecuación (14)

Por último se calculan los pesos w , considerando que la variable latente VL_j satura en las variables observadas $X_{j1}, X_{j2} \dots X_{jk}$, siguiendo el siguiente criterio:

$$\tilde{w}_{jk} = ((\widehat{VL}_j^*)'(\widehat{VL}_j^*))^{-1}(\widehat{VL}_j^*)'X_{jk}$$

Estos pasos se repiten hasta lograr convergencia la cual, por lo general, se determina en función de la diferencia que existe en el valor de w en el paso S en relación al paso $S - 1$. Sanchez propone establecer como criterio de convergencia $|w_{jk}^{S-1} - w_{jk}^S| < 10^{-5}$.

Al finalizar este proceso iterativo, se pasa a las siguientes etapas que consisten en estimar los β 's, γ 's y λ 's que forman parte tanto del modelo estructural, como del de medida.

Las estimaciones de β se obtienen por mínimos cuadrados ordinarios (MCO), tal como se establece a continuación:

$$\hat{\beta}_{ji} = ((\widehat{VL}_i)'(\widehat{VL}_i))^{-1}(\widehat{VL}_i)'(\widehat{VL}_j)$$

Los $\hat{\gamma}$ se estiman de manera análoga.

Por último, los λ 's se estiman como la correlación entre la variable latente y las indicadoras con las que se relaciona, es decir:

$$\hat{\lambda}_{jk} = cor(X_{jk}, \widehat{VL}_j)$$

3. Aplicación

En esta sección se presentan, en forma resumida, los datos con los que se trabajó, describiendo el diseño muestral empleado y el cuestionario utilizado.

3.1. Diseño muestral

La aplicación que se presentará en este trabajo fue realizada sobre los datos obtenidos mediante la aplicación de un cuestionario sobre una muestra probabilística a estudiantes de los cursos superiores de la FCCEEyA., en el año 2009. En esta sección se presentan las principales características del diseño muestral utilizado.

La muestra fue seleccionada en base a un marco muestral que se construyó a partir de las inscripciones a cursos de FCCEEyA en 2009. El diseño muestral usado fue estratificado por conglomerados en 2 etapas y presentó las siguientes características: en una primera instancia se formaron 6 estratos (en base a una clasificación desarrollada en conjunto por investigadores de la cátedra de Metodología de Investigación y del Instituto de Estadística, FCCEEyA - UdelaR (IESTA)) que corresponden aproximadamente a cada uno de los 5 años en los que podía estar cada estudiante en el 2009. Adicionalmente, se propone un 6to estrato para un grupo reducido de materias que corresponden únicamente a la Licenciatura en Administración. Una vez conformados los estratos, se determina que la muestra total se repartirá en forma proporcional a la matrícula de cada estrato.

Al tener definidas las unidades de muestreo, se procede a seleccionar la muestra, proceso que presentó las siguientes etapas:

1. Se sortean los grupos prácticos de cada materia en cada estrato con probabilidad proporcional a la matrícula de cada grupo (conglomerado)
2. Mediante muestreo aleatorio simple (MAS), se seleccionan la misma cantidad de estudiantes en cada grupo seleccionado en la primera etapa. La cantidad de estudiantes de cada grupo es la misma en los 6 estratos.

De esta manera se tiene un diseño muestral que presenta varias ventajas: por un lado, es muy sencillo de implementar en la práctica, ya que no se debe controlar un número diferente de unidades finales de muestreo (estudiantes) en cada grupo y estrato y, por otro, los expansores que se deben usar para las estimaciones son aproximadamente constantes en los 6 estratos, haciendo más sencillos los cálculos.

La muestra finalmente queda conformada por estudiantes que provienen de 60 grupos prácticos (repartidos en forma proporcional en los 6 estratos). Se sortean 12 estudiantes por grupo, lo que determina un tamaño de muestra de 720 estudiantes.

El siguiente cuadro (caudro 1) muestra como quedan repartidos los 60 grupos prácticos en los 6 estratos.

Estrato	1	2	3	4	5	6	Total
# grupos prácticos	21	15	9	9	4	2	60

Cuadro 1: Cantidad de grupos prácticos por estrato

Con la muestra seleccionada, se procedió a realizar el relevamiento de los datos el cual culminó con 647 encuestas realizadas, dejando en evidencia que no fue posible acceder a los 720 estudiantes originalmente estipulados, quedando determinada entonces una tasa de cobertura de la muestra de $647/720 = 90\%$.

En función de esto, al momento de calcular los expansores, lo primero que se hace es analizar el 10% de estudiantes que quedó sin encuestar, con el objetivo de evaluar si se podía pensar que éstos eran una muestra aleatoria de los 720 estudiantes originales, descartando de esta manera un sesgo de selección. Considerando como variables fundamentales el estrato, la edad y el sexo de los estudiantes, se constató que éstas no estaban asociadas a ese 10% que quedó sin encuestar, es decir que ninguno de esos 3 atributos estaban sub o sobre representados. Otros dos aspectos a tener en cuenta, previo al cálculo de los expansores, son los siguientes: por un lado se debe tener en cuenta la existencia de multiplicidad en el marco muestral debido a que hay un número diferente de matrículas correspondientes a cada estudiante, lo que impacta en la probabilidad de selección ya que la unidades primarias de muestreo son conglomerados de matrículas y no de estudiantes, es decir, hay estudiantes que están repetidos y pueden ser encontrados en más de una materia. Por último, debe ser tenido en cuenta el hecho de que la distribución por sexo y

edad presente en la muestra definitiva no es la distribución poblacional, lo cual genera la necesidad de aplicar un proceso de calibración mediante pos-estratificación.

3.2. Cuestionario utilizado

El cuestionario, aplicado sobre la muestra seleccionada, a partir del cual se obtuvieron los datos que resultan el insumo fundamental para el trabajo aquí presentado, resulta de una adaptación del cuestionario utilizado por los investigadores Alves y Raposo de la Universidad de Beira Interior (Portugal). Éste presenta la siguiente estructura: un primer bloque, claramente diferenciado de los demás, que contiene algunas variables de carácter sociodemográfico, como sexo, edad y algunas otras variables que caracterizan al estudiante dentro del ámbito de la facultad, como año de ingreso, año y cantidad de materias en curso, entre otras. Los restantes 8 bloques de preguntas (presentados como bloque A hasta bloque H) presentan todos la misma estructura, se plantea una pregunta general que determina la esencia del bloque y a partir de ella, se establecen una serie de afirmaciones sobre las cuales el estudiante deberá expresar su posición, utilizando una escala Likert que toma valores en el intervalo [1 - 10], donde 1 indicará la mayor discrepancia con lo planteado en la pregunta y 10 el mayor acuerdo.

Los bloques A a H presentan las siguientes características:

- Bloque A - Contiene 12 afirmaciones referentes a las *expectativas* de los estudiantes, previo ingreso a facultad.
- Bloque B - Consta de 6 afirmaciones vinculadas a la *imagen* que tienen los estudiantes sobre la facultad.
- Bloque C - Conformado por 9 afirmaciones asociadas a la *calidad* del servicio que brinda la facultad.
- Bloque D - Contiene 9 afirmaciones asociadas a la *calidad* de los servicios que brinda la facultad con respecto a la biblioteca, beca y cafetería, entre otros.
- Bloque E - Conformado por las mismas 9 afirmaciones que el bloque C, pero asociadas a necesidades/deseos actuales
- Bloque F - Presenta 7 afirmaciones que indagan sobre el *valor percibido*.
- Bloque G - Contiene 6 afirmaciones que refieren a la *satisfacción* de los estudiantes con la facultad.

- Bloque H - Conformado por 5 preguntas que pueden dividirse en 2 subgrupos, las 3 primeras referentes a la *lealtad* de los estudiantes con la facultad, y las 2 últimas asociadas al *boca a boca* que se genera entre los estudiantes.

En este trabajo los bloques D y E no serán considerados y se presentarán resultados para un modelo con 22 de las 45 variables restantes.

4. Resultados

Para la estimación de los modelos en este trabajo solamente se presentan en detalle los resultados al aplicar MCP, método que fuera detallado en la sección 2.2.5 y luego se comparan con los que se obtienen al aplicar MV y MCG.

Los aspectos computacionales de la aplicación presentada en este trabajo fueron realizados con el software libre R - project (R Core Team, 2013). Se utilizaron, en particular, las librerías *plspm* (Sanchez, 2012) y *sem* (Fox, 2006), (Fox et al., 2013).

A continuación se presentan los resultados obtenidos para un modelo cuyo componente estructural propone 15 relaciones entre 7 constructos no observables, mientras que el sub-modelo de medida relaciona estos 7 constructos con 22 variables observadas.

4.1. Modelo estructural

El modelo estructural teórico a ser estimado queda representado gráficamente a través del diagrama presentado en la figura 5.

Este modelo es estimado por mínimos cuadrados parciales. En el cuadro 2 se exponen las estimaciones puntuales (Original) para cada uno de los β 's y γ 's presentes en el modelo. Para ser validadas, estas estimaciones serán comparadas contra las obtenidas para 100 muestras *bootstrap*, reportando la media muestral, el error estándar, y los cuantiles empíricos para cada parámetro.

A partir de los resultados presentados en el cuadro 2 se concluye que de las 15 relaciones propuestas en el modelo estructural, hay 2 que no se confirman. Éstas refieren a la influencia que tienen las *expectativas* sobre la construcción de los conceptos de *valor percibido* y de *satisfacción*.

El modelo estructural estimado bajo MCP (cuya representación teórica se muestra gráficamente en la figura 5) luce de la siguiente manera (figura 6):

Modelo Teórico Propuesto

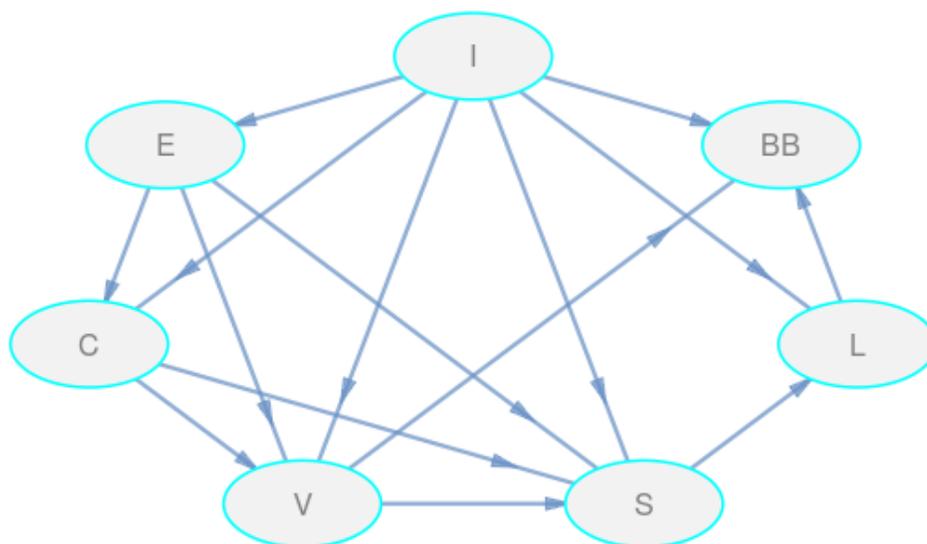


Figura 5: Modelo estructural

Relación	Original	\bar{x} .Boot	E.E.	$q(0,025)$	$q(0,975)$
I → E	0.61	0.62	0.04	0.55	0.68
I → C	0.53	0.53	0.03	0.46	0.59
I → V	0.42	0.41	0.05	0.32	0.52
I → S	0.23	0.22	0.04	0.15	0.31
I → L	0.16	0.15	0.05	0.06	0.25
I → BB	0.23	0.23	0.04	0.13	0.32
E → C	0.40	0.40	0.04	0.34	0.47
E → V	0.08	0.08	0.05	-0.02	0.18
E → S	-0.00	-0.01	0.04	-0.09	0.07
C → V	0.31	0.32	0.06	0.22	0.43
C → S	0.30	0.30	0.06	0.19	0.41
V → S	0.39	0.39	0.04	0.32	0.47
V → BB	0.30	0.30	0.05	0.21	0.39
S → L	0.62	0.63	0.04	0.54	0.70
L → BB	0.38	0.38	0.04	0.28	0.47

Cuadro 2: Modelo estructural - Estimaciones

Modelo Estructural Estimado

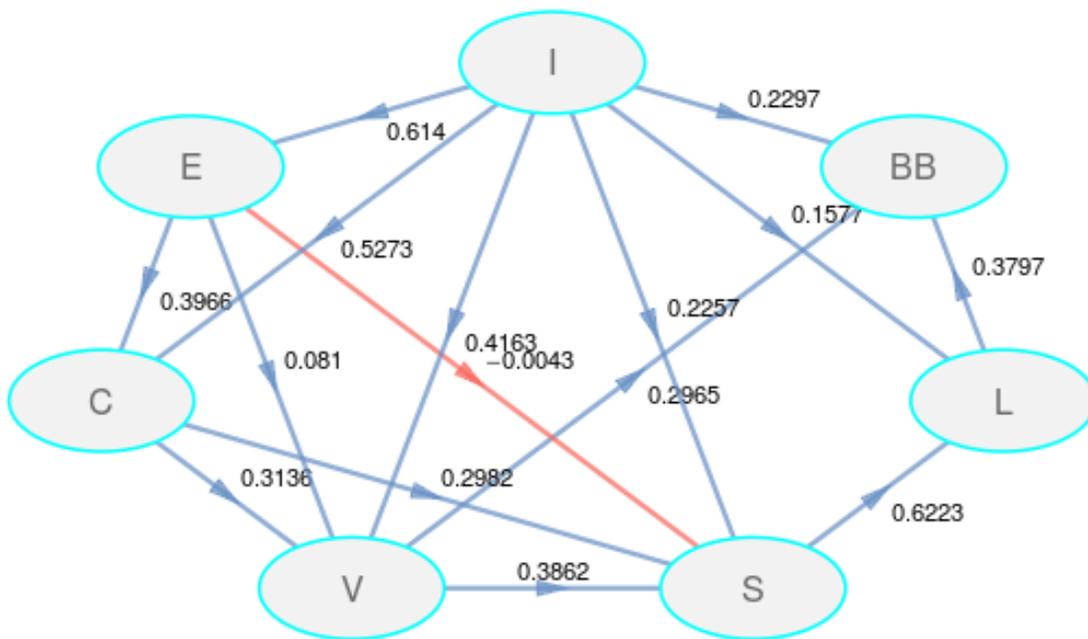


Figura 6: Modelo estructural estimado

Otra medida que resulta de interés al momento de evaluar el modelo estructural es la proporción de varianza de cada variable latente que logra ser explicada por las variables, también latentes, que la causan directamente. Esta proporción queda determinada a partir del coeficiente de determinación R^2 .

Variable	R^2
E	0.38
C	0.69
S	0.67
VP	0.56
L	0.55
BB	0.63

Cuadro 3: R^2 para el modelo estructural

A partir del cuadro 3 puede concluirse que *calidad* es el constructo que mejor queda explicado por las variables que lo preceden, las que logran explicar un 69 % de la variabilidad total del constructo. En segundo lugar se encuentra aquella variable que refleja el grado de *satisfacción* de los estudiantes, el 67 % de la varianza total de este constructo queda explicado por las variables *expectativas*, *calidad*, *valor percibido e imagen*. En el otro extremo, se encuentran las *expectativas*, las cuales se entiende quedan determinadas por la *imagen* que tienen los estudiantes sobre la facultad, sin embargo ésta parece no resultar suficiente, ya que sólo logra captar un 38 % de la variabilidad total.

Además del análisis de los efectos *directos* que tiene una variable latente sobre otra u otras, se pueden analizar las relaciones *indirectas* que se generan dentro del modelo estructural.

Tal como se observa en la figura 7 el constructo que tiene mayor impacto indirecto sobre los demás es aquel que refiere a la *imagen* que tienen los estudiantes sobre la facultad. En este sentido cabe resaltar la relación que se da entre este constructo y los que miden la *satisfacción*, la *lealtad* y el *boca a boca* donde el efecto indirecto es tres veces mayor que el directo.

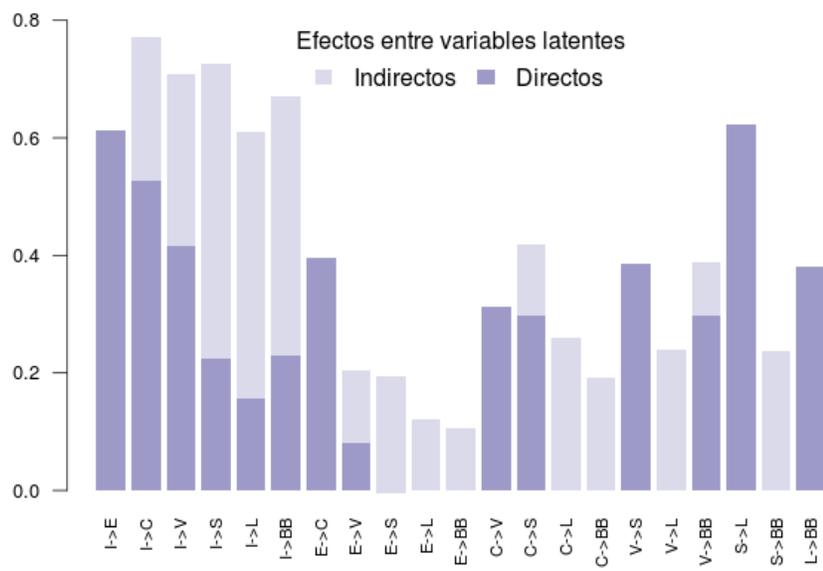


Figura 7: Modelo estructural - Efectos directos e indirectos

4.2. Modelo de medida

El modelo de medida estimado por mínimos cuadrados parciales puede verse gráfica y numéricamente en la figura 8 y cuadro 4, respectivamente.

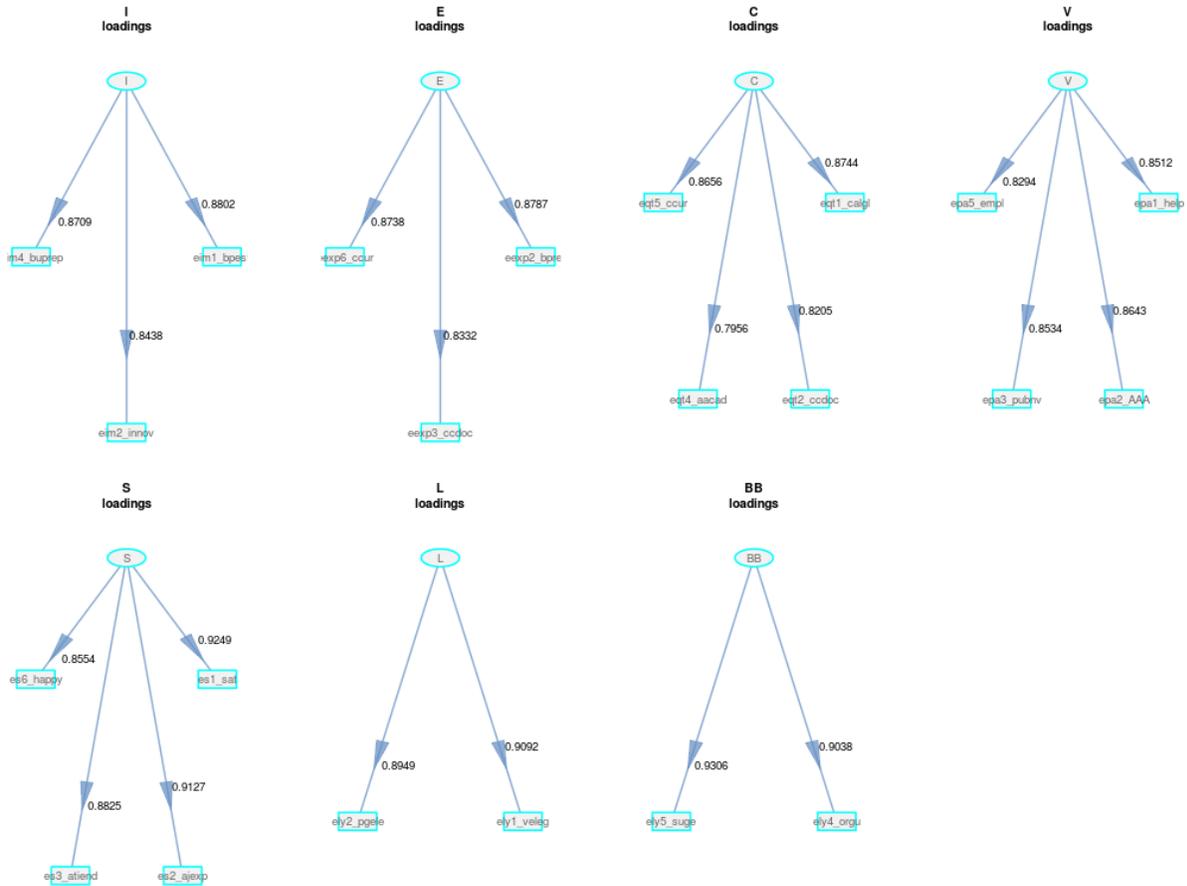


Figura 8: Modelo de medida - Estimado

Al observar las estimaciones obtenidas para cada uno de los λ 's presentes en el submodelo de medida, lo primero que cabe destacar es que, tal como puede verse en el cuadro 4, los parámetros estimados resultan todos mayores a 0.7 (valor mínimo aceptado para validar estas estimaciones).

Por otra parte, se debe recordar que este modelo se encuentra en el ámbito del análisis factorial, donde el concepto de comunalidad, presentado en la sección 2.1, resulta fundamental. Este concepto refiere a la proporción de varianza original de cada variable

V.Obs	V.Lat.	λ	Comunalidad
eexp2_bprep	E	0.88	0.77
eexp3_ccdoc	E	0.83	0.69
eexp6_ccur	E	0.87	0.76
eim1_bpest	I	0.88	0.77
eim2_innov	I	0.84	0.71
eim4_buprep	I	0.87	0.76
ely1_veleg	L	0.91	0.83
ely2_pgele	L	0.89	0.80
ely4_orgu	BB	0.90	0.82
ely5_suge	BB	0.93	0.87
epa1_help	V	0.85	0.72
epa2_AAA	V	0.86	0.75
epa3_pubnv	V	0.85	0.73
epa5_empl	V	0.83	0.69
eqt1_calgl	C	0.87	0.76
eqt2_ccdoc	C	0.82	0.67
eqt4_aacad	C	0.80	0.63
eqt5_ccur	C	0.87	0.75
es1_sat	S	0.92	0.86
es2_ajexp	S	0.91	0.83
es3_atiend	S	0.88	0.78
es6_happy	S	0.86	0.73

Cuadro 4: Modelo de medida - Estimado

observada, que queda explicada por el factor que satura sobre ella.

En el cuadro 4 se observa que para el bloque que contiene las variables referentes a las *expectativas* que tienen los estudiantes sobre la facultad, se tiene que el factor común logra explicar más del 69 % de la varianza original de cada una de las 3 variables.

En cuanto a las variables que determinan la *imagen* que tienen los estudiantes de la FC-CEEyA sobre ésta, puede verse que las comunalidades toman valores entre 0.71 y 0.77. La variable que mejor queda explicada por el constructo *imagen* es aquella que refiere a la visión general que tienen los estudiantes sobre la Universidad, como lugar donde estudiar (*eim1_bpest*).

Al considerar las variables que conforman el bloque referente a la evaluación que hacen los estudiantes sobre la *calidad* del servicio brindado por la facultad se muestra que la variable observada que mejor queda explicada por la variable latente *calidad* es la que refiere a la calidad global de enseñanza (*eqt1_calgl*), ya que el 76 % de su varianza original es captada por el constructo. En el otro extremo, se encuentra la afirmación que hace referencia al ambiente académico (*eqt4_aacad*), donde el factor logra explicar el 63 % de la varianza original.

En lo que refiere al bloque de preguntas específicas sobre *satisfacción*, puede verse que este constructo logra captar una cantidad importante (entre un 73 % y un 86 %) de la varianza original de las variables observadas que lo conforman.

Considerando aquellas variables que son entendidas como causantes de la *satisfacción*, sólo resta presentar el bloque de preguntas que refiere al *valor percibido*. Tal como puede verse en el cuadro 4, las cuatro variables observadas, aquí consideradas, que conforman el constructo *valor percibido* quedan bien explicadas por éste (comunalidad mayor a 0.68).

En cuanto a las variables consideradas como consecuencias de la satisfacción, en el cuadro 4, se presentan aquellas que conforman el constructo *lealtad*, donde se ve que la variable que mejor queda explicada por éste es “*Si tuviera que decidir nuevamente, volvería a elegir esta facultad*” (*ely1_veleg*).

Por último, puede verse que la varianza de aquellas variables que forman el constructo *boca a boca* resulta captada en gran proporción por este factor, sobre todo para la variable “*Recomendaría esta facultad a un amigo*” (*ely5_suge*).

5. Discusión

A los efectos de poder tener comparabilidad con los resultados obtenidos por otros investigadores (Alves y Raposo) (Alves and Raposo, 2005), se estima mediante MV, que supone distribución multinormal. Ésta fue testeada a través de los estadísticos de simetría y kurtosis propuestos por Mardia (Kankainen et al., 2004), a partir de los cuales se rechazó la hipótesis de existencia de normalidad multivariada. Este resultado es el esperado considerando que las variables observadas son variables discretas que toman valores en el intervalo [1 - 10], por lo que la normalidad nunca podría resultar más que una aproximación.

Dejando de lado el supuesto de normalidad se estima, también, a través de MCG y se comparan los resultados obtenidos a partir de los 3 métodos.

Parámetro	MCP	$q(0,025)$ - $q(0,095)$	MV	$\Pr(> z)$	MCG	$\Pr(> z)$
$\gamma_{I \rightarrow E}$	0.61	[0.55;0.67]	0.63	0.00	0.07	0.14
$\gamma_{I \rightarrow C}$	0.53	[0.47;0.58]	0.48	0.00	0.27	0.00
$\gamma_{I \rightarrow V}$	0.42	[0.32;0.49]	0.45	0.00	0.16	0.01
$\gamma_{I \rightarrow S}$	0.23	[0.16;0.30]	0.23	0.00	0.05	0.18
$\gamma_{I \rightarrow L}$	0.16	[0.06;0.26]	0.13	0.00	0.04	0.53
$\gamma_{I \rightarrow BB}$	0.23	[0.14;0.29]	0.10	0.04	0.07	0.15
$\beta_{E \rightarrow C}$	0.40	[0.33;0.47]	0.34	0.00	0.05	0.30
$\beta_{E \rightarrow V}$	0.08	[0.00;0.16]	0.03	0.52	0.01	0.85
$\beta_{E \rightarrow S}$	-0.00	[-0.07;0.08]	-0.04	0.42	-0.05	0.12
$\beta_{C \rightarrow V}$	0.31	[0.21;0.42]	0.36	0.00	0.06	0.17
$\beta_{C \rightarrow S}$	0.30	[0.20;0.40]	0.30	0.00	0.05	0.06
$\beta_{V \rightarrow S}$	0.39	[0.31;0.46]	0.36	0.00	0.11	0.01
$\beta_{V \rightarrow BB}$	0.30	[0.22;0.39]	0.21	0.00	0.18	0.00
$\beta_{S \rightarrow L}$	0.62	[0.53;0.70]	0.76	0.00	-0.04	0.52
$\beta_{L \rightarrow BB}$	0.38	[0.30;0.45]	0.60	0.00	0.30	0.00

Cuadro 5: Modelo estructural - Comparación de MCP, MV y MCG

Si se observan los resultados del cuadro 5 se ve que para el modelo estructural estimado no existen diferencias relevantes entre las estimaciones con MCP y MV, donde se encuentra que las mismas variables no resultan significativas en ambos casos. La relación $E \rightarrow V$ y $E \rightarrow S$ coinciden en no ser significativas con estimaciones cercanas a 0, intervalo bootstrap que contiene al 0 y un p-valor mayor a $\alpha = 0,05$.

Al comparar los resultados contra los obtenidos por MCG, se observan mayores diferencias: 10 de las 15 relaciones propuestas, no logran ser confirmadas. El constructo *expectativas* parece quedar fuera del modelo, ya que la variable que se propone como causa (*imagen*)

parece no serlo y, a la vez, ni la *calidad*, ni el *valor percibido*, ni la *satisfacción* parecen ser causadas por ésta.

En el cuadro 6 se observa que para el modelo de medida tampoco existen diferencias entre las estimaciones obtenidas por MCP y MV. En el caso del primer método, todas las estimaciones superan el valor mínimo *aceptable* (0.7) y para el segundo todas las variables resultan significativas al 5% (asumiendo normalidad). Al comparar los resultados con los obtenidos por MCG, sí se observan diferencias: existen 3 parámetros para los cuales no se tiene suficiente información estadística que permita rechazar la hipótesis de que estos sean nulos. Éstos refieren a 2 variables observadas que relacionadas con el constructo *calidad* (*eqt4_aacad*: ambiente académico y *eqt5_ccur*: contenido de los cursos) y 1 variable observada referente al factor *satisfacción* (*es6_happy*: grado de felicidad por haber elegido esta facultad).

Parámetro	MCP	MV	Pr(> z)	MCG	Pr(> z)
$\lambda_{exp2_bprep \rightarrow E}$	0.88	0.83	0.00	0.51	0.00
$\lambda_{exp3_ccdoc \rightarrow E}$	0.83	0.79		0.94	
$\lambda_{exp6_ccur \rightarrow E}$	0.87	0.81	0.00	0.53	0.00
$\lambda_{eim1_bpest \rightarrow I}$	0.88	0.84	0.00	0.21	0.00
$\lambda_{eim2_innov \rightarrow I}$	0.84	0.72	0.00	0.28	0.00
$\lambda_{eim4_buprep \rightarrow I}$	0.87	0.72		0.96	
$\lambda_{ely1_veleg \rightarrow L}$	0.91	0.86		0.84	
$\lambda_{ely2_pgele \rightarrow L}$	0.89	0.75	0.00	0.26	0.00
$\lambda_{ely4_orgu \rightarrow BB}$	0.90	0.72	0.00	0.37	0.00
$\lambda_{ely5_suge \rightarrow BB}$	0.93	0.99		0.99	
$\lambda_{epa1_help \rightarrow V}$	0.85	0.82		0.76	
$\lambda_{epa2_AAA \rightarrow V}$	0.86	0.83	0.00	0.58	0.00
$\lambda_{epa3_pubnv \rightarrow V}$	0.85	0.80	0.00	0.41	0.00
$\lambda_{epa5_empl \rightarrow V}$	0.83	0.76	0.00	0.55	0.00
$\lambda_{eqt1_calgl \rightarrow C}$	0.87	0.91		1.33	
$\lambda_{eqt2_ccdoc \rightarrow C}$	0.82	0.76	0.00	0.16	0.01
$\lambda_{eqt4_aacad \rightarrow C}$	0.80	0.72	0.00	0.04	0.36
$\lambda_{eqt5_ccur \rightarrow C}$	0.87	0.83	0.00	0.09	0.06
$\lambda_{es1_sat \rightarrow S}$	0.92	0.93		1.33	
$\lambda_{es2_ajexp \rightarrow S}$	0.91	0.90	0.00	0.29	0.00
$\lambda_{es3_atiend \rightarrow S}$	0.88	0.84	0.00	0.16	0.00
$\lambda_{es6_happy \rightarrow S}$	0.86	0.83	0.00	0.01	0.87

Cuadro 6: Modelo de medida - Comparación de MCP, MV y MCG

Índice de satisfacción estudiantil - FCCEEyA

Considerando las estimaciones obtenidas a través del método de MV se puede calcular un índice de satisfacción estudiantil para la FCCEEyA, a partir de la fórmula propuesta por Fornell et al. (Claes et al., 1996). Esto se hace trabajando con las estimaciones no estandarizadas de los factores de carga que vinculan el factor *satisfacción* con cuatro de las variables observadas que conforman este constructo, con el siguiente algoritmo:

$$IS = \frac{\sum_i w_i \bar{y}_i - \sum_i w_i}{10 \sum_i w_i - \sum_i w_i} * 100 = \frac{\sum_i w_i \bar{y}_i - \sum_i w_i}{9 \sum_i w_i} * 100 \quad (15)$$

Variable Medida (y_i)	Estimación no estandarizada Uruguay (w_i)	Media variable medida (\bar{y}_i)	$w_i \bar{y}_i$
S_1	1.00	7.45	7.45
S_2	0.97	7.03	6.81
S_3	0.89	6.70	5.96
S_6	0.98	7.53	7.37
<i>Total</i>	3.84	-	27.59

Cuadro 7: Índice de satisfacción estudiantil FCCEEyA - modelo UP2

A partir de lo expuesto en el cuadro 7, el IS para la FCCEEyA es 69%.

6. Conclusiones

En cuanto al objetivo principal de este trabajo, éste apuntaba fundamentalmente a la evaluación de un instrumento de medida para evaluar el nivel de satisfacción estudiantil para los cursos de educación superior de la FCCEEyA. Ésto implicaba llevar a cabo la modelización de la satisfacción a través de la aplicación de modelos de ecuaciones estructurales, a partir de los cuales se generan nuevos objetivos.

En lo que refiere al modelo estructural, el objetivo perseguido al plantearlo era ver si determinadas relaciones, tomadas tanto del ECSI como de las investigaciones portuguesas (Alves and Raposo, 2005), se confirmaban para el caso de la FCCEEyA. De esto surgen conclusiones que apuntan por un lado, a la comparación directa con, por ejemplo, la Universidad de Beira Interior, las cuales establecen que existen diferencias en cómo se elabora el constructo *satisfacción* en ambos casos (Vernazza, 2013).

En cuanto a los diseños muestrales utilizados, para el caso de la FCCEEyA se usó un diseño muestral complejo, el cual determinaba el manejo de pesos autoponderados, lo

cual finalmente no resultó posible, ya que existió la necesidad de calibrar y trabajar con multiplicidad. Para el caso de Portugal (Alves and Raposo, 2005), el diseño muestral, proponía tener un número fijo de estudiantes por Universidad (250) pero esto no resulta (para 2 Universidades el tamaño se redujo sensiblemente).

De todas formas, si bien originalmente ambos diseños difieren, los resultados presentados en ambos casos se obtienen bajo el supuesto de un muestreo aleatorio simple.

Para este trabajo se hace una depuración de las 45 variables originales, considerando solamente 22, 18 de las cuales coinciden con las utilizadas por los investigadores portugueses (Alves and Raposo, 2005), y 4 más que surgen de tener en cuenta la consistencia interna dentro de cada bloque (Vernazza, 2013).

Los resultados obtenidos en este trabajo, presentados en secciones previas, se obtuvieron mediante MCP pero a los efectos de tener comparabilidad con los resultados de los investigadores portugueses (Alves and Raposo, 2005), se estiman con el método MV que tiene el supuesto de distribución multinormal de las variables observadas, aspecto que se analizó y se corroboró que no se cumplía. La violación de este supuesto, fundamental dentro del análisis factorial, influye sobre las estimaciones de los parámetros involucrados en el modelo, afectando directamente las decisiones que se tomen a partir de éstas. De todas formas, en este trabajo para la estimación bajo MV se asume que sí existe multinormalidad de las variables y se presentan resultados, y por ende conclusiones, respaldados en este supuesto.

Con las estimaciones obtenidas bajo MV se puede calcular el Índice de Satisfacción, que presenta un valor de 0.69.

Además de los resultados obtenidos por MV y MCP, se presentan estimaciones bajo MCG. Los resultados presentados a partir de la utilización de este último método son los que presentan mayores diferencias con respecto a las obtenidas para los otros dos métodos.

6.1. Consideraciones a futuro

Como consideraciones para futuras investigaciones se sugiere ver cómo impacta en los estadísticos de prueba de bondad de ajuste para MV, la violación del supuesto de multinormalidad, ya que la mayoría de estos estadísticos, son reglas empíricas del estilo “rule of thumb” basados en supuestos que aquí no se verifican. Si la intención resulta mantenerse dentro de la máxima verosimilitud, se puede considerar, por ejemplo, la propuesta de Oberski y Satorra (Oberski y Satorra, 2013), quienes sugieren un proceso de estimación en dos etapas e incorporar componentes extra de varianza (Satorra y Muthen, 1995), (Stapleton, 2006), (Stapleton, 2008).

En cuanto al método de MCP es importante tratar de incorporar, en las rutinas de estimación, el peso diferente que tienen las respuestas, al ser una muestra probabilística compleja con expansores diferentes. Actualmente el software usado, no admite forma de ponderar los datos, aspecto que sí se podía hacer para el método MV que puede trabajar con una matriz de varianzas y covarianzas expandida.

En cuanto al cálculo del índice de satisfacción es necesario evaluar como sería la adaptación al trabajar con factores de carga estandarizados.

Referencias

- Alves, H. y Raposo, M. (2005). La medición de la satisfacción en la enseñanza universitaria: El ejemplo de la Universidade da Beira Interior. *International Review on Public and Nonprofit Marketing*, 1(1):73–88.
- Blanco, J. (2006). *Introducción al análisis multivariado*. Universidad de la República.
- Blanco, R. J. y Blanco Peck, R. (2007). La medición de la calidad de servicios en la educación universitaria. *Cuaderno de Investigación en la Educación*, 22(7):121–136.
- Casas Guillén, M. (2008). Los modelos de ecuaciones estructurales y su aplicación en el índice europeo de satisfacción del cliente. Technical report, Facultad de Económicas, Universidad San Pablo, CEU.
- Claes, F., Michael D., J., Eugene W., A., Jaesung, C., y Bryant, B. E. (1996). The american customer satisfaction index: Nature, purpose, and findings. *Journal of Marketing*, 60:7–18.
- Dillon, W. y Goldstein, M. (1984). *Multivariate analysis methods and applications*. Wiley.
- Fox, J. (2006). Structural equation modeling with the package in r. *STRUCTURAL EQUATION MODELING*, 13(3):465–486.
- Fox, J., Nie, Z., y Byrnes, J. (2013). *sem: Structural Equation Models*. R package version 3.1-1.
- Kankainen, A., Taskinen, S., y Oja, H. (2004). On mardia’s tests of multinormality. In Hubert, M., Pison, G., Struyf, A., t Van Aelst, S., editors, *Theory and Applications of Recent Robust Methods*, Statistics for Industry and Technology, pages 153–164. Birkhäuser Basel.
- Kline, R. (2011). *Principles and Practice of Structural Equation Modeling*. The Guilford Press.
- Oberski, D. y Satorra, A. (2013). Measurement error models with uncertainty about the error variance. *Structural Equation Modeling*, 20:409–428.
- Peña, D. (2002). *Análisis de datos multivariantes*. McGraw-Hill/Interamericana, Madrid.
- R Core Team (2013). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
- Revelle, W. (2004). *An introduction to psychometric theory with applications in R*. Lawrence Erlbaum Associates, Mahwah, N.J.

- Sanchez, G. (2012). *PLS Path Modeling with R*.
- Satorra, A. y Muthen, B. (1995). Complex sample data in structural equation modeling. *Sociological methodology*, 25:267–316.
- Stapleton, L. (2006). An assessment of practical solutions for structural equation modeling with complex sample data. *Structural Equation Modeling*, 13(1):28–58.
- Stapleton, L. (2008). Variance estimation using replication methods in structural equation modeling with complex sample data. *Structural Equation Modeling*, 15(2):183–210.
- Vernazza, E. (2013). Evaluación de un instrumento de medición del nivel de satisfacción estudiantil en los cursos de formación superior de la FCCEEyA de la UDELAR a través de la aplicación de Structural Equation Modelling (SEM). Informe de pasantía, Fac. de CCEEyA - UdelaR.