

---

## **Efecto de los métodos de estimación en las modelaciones de estructuras de covarianzas sobre un modelo estructural de evaluación del servicio de clases**

**Effect of estimation methods of structural equations models on a structural model to evaluate the class service**

Carlos Eduardo Valdivieso Taborga<sup>a</sup>  
cvaldivieso@upb.edu

---

### **Resumen**

El propósito de esta investigación ha sido realizar un análisis comparativo de los métodos de estimación de parámetros de las modelaciones de estructuras de covarianzas, con el fin de determinar el método más adecuado, usando un modelo estructural de evaluación del servicio de clases, que consiste en las relaciones entre la calidad de este servicio y el valor percibido sobre las intenciones de comportamiento de posclase, actuando como variable mediadora la satisfacción estudiantil.

La metodología para establecer el método más adecuado para la estimación se basó en aspectos limitantes del método considerado como la suposición de normalidad y tamaño muestral y en aspectos de la superioridad del modelo como el poder predictivo y los índices de bondad de ajuste.

Se ha establecido que el mejor método de estimación de parámetros del modelo estructural de evaluación del servicio de clases investigado es el de mínimos cuadrados no ponderados, a causa de que: 1) se comprobó que los datos no cumplen la suposición de normalidad multivariante, 2) es el único método en el que el modelo presenta índices de bondad de ajuste adecuados, y 3) el modelo tiene un alto poder predictivo. En general se ha observado que existen desviaciones porcentuales de los valores de los paths del modelo usando los distintos métodos comparados con el ULS, que van desde -45 % a 47 %.

Utilizando el método ULS se ha establecido que la satisfacción estudiantil es un mediador poderoso entre la calidad de servicio de clases y el valor percibido, sobre las intenciones de comportamiento de posclase.

---

<sup>a</sup>Docente investigador. Departamento de Ciencias Exactas. Universidad Privada Boliviana (UPB).

**Palabras clave:** máxima verosimilitud, mínimos cuadrados no ponderados, servicio de clases, modelo de estructura de covarianza.

### Abstract

The purpose of this research was to perform a comparative analysis of methods for estimating parameters of MEC, in order to determine the most appropriate method, specifically using a structural model of class service evaluation. This model consists of the relationships between perceived service quality and perceived value classes on behavioral intentions of postclass, acting as a mediator variable student satisfaction.

The methodology for establishing the most appropriate method is based on the limiting factors of the method considered (assumption of normality and sample size) and aspects of the superiority of the model (predictive power and goodness of fit indices). It has been established that the best method for estimating parameters of an structural model for service class evaluation is the unweighted least squares (ULS), because: 1) it was found that the data did not meet the assumption of multivariate normality, 2) is the only method in which the model presents indices appropriate fit, and 3) the model has high predictive power. In general it has been observed that there are percentage deviations from the values of the paths of the model using the various methods compared with the ULS ranging from -45 % to 47 %.

Using the ULS method has been established that student satisfaction is a powerful mediator between service quality and perceived value classes on behavioral intentions of postclass.

Using this method, it was established that student satisfaction is a powerful mediator between class service quality and perceived value on after-class behavioral intentions.

**Key words:** maximum likelihood, unweighted least squares, class service, covariate structure models.

## 1. Introducción

Cuando se lleva a cabo una modelación de estructuras de covarianzas (MEC), el investigador debe decidir sobre varios aspectos: la especificación del modelo, su identificación, el método de estimación de parámetros que utilizará y el procedimiento para evaluar el ajuste del modelo (Bollen 1989). El presente trabajo se centra en realizar un análisis comparativo de los diferentes métodos de estimación de parámetros, con el fin de decidir cuál es el más adecuado, dependiendo del cumplimiento de algunos aspectos limitantes de cada método. Si las variables que entran en el modelo estructural siguen una distribución normal multivariante, suele utilizarse el método de máxima verosimilitud (ML, por sus siglas en inglés) por sus propiedades asintóticas (proporciona estimadores insesgados, consistentes,

eficientes y normalmente distribuidos). Cuando el supuesto de normalidad no se cumple, uno de los métodos más empleados es el de mínimos cuadrados no ponderados (ULS, por sus siglas en inglés), ya que no requiere de muestras grandes. Otros métodos alternativos son el de mínimos cuadrados generalizados (GLS, por sus siglas en inglés) y el de mínimos cuadrados ponderados (WLS o ADF, por sus siglas en inglés). Estos métodos difieren en sus supuestos y en su función de ajuste, y proporcionan estimaciones de los parámetros ligeramente diferentes (Ximénez & García 2005), por lo cual, su elección es de primordial importancia en el estudio de modelos en cualquier área de investigación.

En la literatura del marketing de servicios varios investigadores han estudiado las relaciones existentes entre la calidad de servicio, satisfacción, valor percibido e intenciones de comportamiento, mediante modelos alternativos (Brady et al. 2005, Choi et al. 2004, Chen 2008). Las principales conclusiones que se pueden extraer después del análisis de los distintos modelos son las siguientes:

- El valor percibido, la calidad de servicio, la satisfacción y las intenciones de comportamiento, deben ser medidos preferentemente mediante un constructo unidimensional (con múltiples ítems) o multidimensional, y no mediante una escala de un solo ítem. La calidad de servicio percibida generalmente es un constructo de segundo orden, operacionalizado mediante el instrumento SERVQUAL (ServiceQuality), desarrollado por Parasuraman et al. (1988).
- La calidad percibida es un antecedente del valor percibido. A su vez, la calidad y el valor son antecedentes de la satisfacción.
- La calidad de servicio, el valor percibido y la satisfacción son poderosos predictores de las intenciones de comportamiento.
- La satisfacción es una consecuencia del valor percibido.
- Las evaluaciones de la calidad de servicio, valor percibido, satisfacción e intenciones de comportamiento, pueden variar entre clientes, culturas y servicios.
- Un entendimiento de cómo se relacionan estos constructos y de cómo medirlos adecuadamente es esencial para entender el comportamiento del consumidor.
- Las investigaciones futuras deben focalizarse en el estudio de modelos comprensivos.

En la presente investigación, para realizar el análisis comparativo de los métodos de estimación de los MEC, se ha elegido un modelo estructural comprensivo de evaluación del servicio de clases, en el cual la calidad de servicio de clases es antecedente del valor percibido del servicio de clases y de la satisfacción estudiantil; el valor percibido es antecedente de la satisfacción estudiantil, y todos ellos son antecedentes de las intenciones de comportamiento de posclase, como se muestra en la Figura 1.

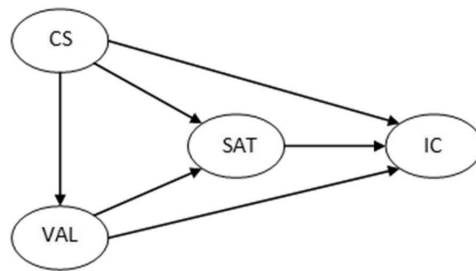


Figura 1: *Modelo comprehensivo de evaluación del servicio* (CS: calidad de servicio; SAT: satisfacción; VAL: valor; e IC: intenciones de comportamiento). Fuente: Modificado de Choi et al. (2004).

Para llevar a cabo este análisis se siguió la siguiente metodología. Primero se recopilaron datos sobre percepciones de los constructos de interés. Luego se realizaron pruebas de idoneidad de los datos recopilados: se verificó que no existan datos faltantes ni atípicos, se realizó una prueba de normalidad multivariante mediante el método del gráfico chi-cuadrado (Johnson & Wichern 1998) y se verificó la adecuación de los datos al patrón poblacional. Mediante un análisis factorial exploratorio (AFE) se estableció una estructura dimensional adecuada y baterías para los constructos de interés. Posteriormente se realizó la modelación MEC con los distintos métodos de estimación de parámetros, comparándose los valores de los paths de cada relación del modelo de evaluación del servicio de clases, el poder predictivo del modelo (dado por el valor de la correlación múltiple cuadrada de las intenciones de comportamiento de posclase) y los índices ad hoc de bondad de ajuste, para luego definir el método más adecuado con base en el que presenta mayor poder predictivo y mejores índices de bondad de ajuste.

## 2. Fundamento teórico

### 2.1. Modelos de estructura de covarianza (MEC)

#### 2.1.1. Introducción

Siguiendo a Long (1983), los MEC intentan explicar las relaciones entre un conjunto de variables observadas, basándose en las relaciones de variables latentes, caracterizando dichas relaciones mediante las covarianzas de las variables observadas.

Todo modelo de estructuras de covarianza se descompone en dos: el primero de ellos es el componente estructural, que incluye las relaciones entre los factores latentes, el segundo es el componente de medida, donde cada factor latente ha de

medirse mediante variables observadas que midan ese factor, ya sea dependiente o independiente. Para modelar la imperfección de las medidas, los MEC incorporan los denominados errores de medida (Uriel & Aldás 1996).

### 2.1.2. Formalización matemática de los MEC

En su forma más general, los MEC se pueden representar a partir de tres sistemas de ecuaciones:

1. Los que recogen la relación causal entre las variables latentes (componente estructural).
2. Los que se ocupan de reflejar el componente de medida; es decir los indicadores de cada factor. Para las variables observadas que miden un factor dependiente ( $Y$ );
3. y para los que miden un factor independiente ( $X$ ).

#### Componente estructural

El componente estructural de los MEC recoge la relación causal entre las variables latentes:

$$\eta = B\eta + \Gamma\xi + \zeta \quad (1)$$

Donde, como se resume en la Tabla 1,  $\eta$  es un vector  $r \times 1$  que contiene los  $r$  factores dependientes,  $B$  es una matriz  $r \times r$  que contiene los coeficientes que relacionan entre sí a los factores dependientes,  $\Gamma$  es una matriz  $r \times s$  con los coeficientes que relacionan los  $r$  factores dependientes con los  $s$  independientes.  $\xi$  es un vector  $s \times 1$  con los factores independientes.  $\zeta$  es un vector  $r \times 1$  que contiene los errores asociados con los factores dependientes.

La ecuación anterior, tras restar  $B\eta$  a ambos miembros:  $\eta - B\eta = \Gamma\xi + \zeta$ , y definir:  $\ddot{B} = I - B$ , se puede expresar de la siguiente forma:  $\ddot{B}\eta = \Gamma\xi + \zeta$ .

Tabla 1: *Resumen del componente estructural de los MEC. Fuente: Long (1983)*

Matriz	Dimensión	Media	Covarianza	Dimensión	Descripción
$\eta$	$r \times 1$	0	$Cov(\eta) = E(\eta\eta')$	$r \times r$	Factores dependientes
$\xi$	$s \times 1$	0	$\Phi = E(\xi\xi')$	$s \times s$	Factores independientes
$\zeta$	$r \times 1$	0	$\Psi = E(\zeta\zeta')$	$r \times r$	Errores en las ecuaciones
$B$	$r \times r$	-	-	-	Efectos directos $\eta$ sobre $\eta$
$\Gamma$	$r \times s$	-	-	-	Efectos directos $\xi$ sobre $\eta$

Las hipótesis con respecto al componente estructural son las siguientes:

1. Las variables están medidas en desviaciones respecto a su media, lo que implica que:  $E(\eta) = E(\xi) = E(\zeta) = 0$ .
2. Los términos de error de los factores dependientes y los factores independientes están incorrelacionados entre sí:  $E(\xi\xi') = E(\zeta\zeta') = 0$ .
3. También se asume que  $\ddot{B}$  es no singular (existe su inversa).

**Componente de medida:** los sistemas de ecuaciones para los componentes de medida son:

$$x = \Lambda_X \xi + \delta \quad y = \Lambda_Y \eta + \varepsilon \quad (2)$$

Los elementos son el vector  $x$  de dimensión  $q \times 1$  que contiene las variables observadas que sirven para medir los factores independientes ( $q$  es el número de esas variables observadas); y es el vector de variables observadas (hay  $p$  de ellas, por ello su dimensión es  $p \times 1$ ) que miden los factores dependientes. Las matrices que contienen los coeficientes de regresión o cargas factoriales entre esas variables y sus respectivos factores vienen denotadas por  $\Lambda_x$  para los factores independientes y  $\Lambda_y$  para los dependientes. Sus dimensiones son respectivamente  $q \times s$  y  $p \times r$ , siendo  $s$  y  $r$  el número de factores de cada tipo. Esta información se resume en la Tabla 2.

Tabla 2: Resumen del componente de medida de los MEC. Fuente: Long (1983)

Matriz	Dimensión	Media	Covarianza	Dimensión	Descripción
$\eta$	$r \times 1$	0	$Cov(\eta) = E(\eta\eta')$	$r \times r$	Factores dependientes
$\xi$	$s \times 1$	0	$\Phi = E(\xi\xi')$	$s \times s$	Factores independientes
$x$	$q \times 1$	0	$\Sigma_{xx} = E(xx')$	$q \times q$	Variables observadas independientes
$y$	$p \times 1$	0	$\Sigma_{yy} = E(yy')$	$p \times p$	Variables observadas dependientes
$\Lambda_x$	$q \times s$	—	—	—	Cargas de $x$ sobre $\xi$
$\Lambda_y$	$p \times r$	—	—	—	Cargas de $y$ sobre $\eta$
$\delta$	$q \times 1$	0	$\Theta_\delta = E(\delta\delta')$	$q \times q$	Errores de $x$
$\varepsilon$	$p \times 1$	0	$\Theta_\varepsilon = E(\varepsilon\varepsilon')$	$p \times p$	Errores de $y$

Las hipótesis que se postulan sobre las variables de las ecuaciones anteriores son:

1. Las variables están medidas en desviaciones sobre la media:  $E(x) = E(\delta) = E(y) = E(\varepsilon) = 0$ .
2. Los factores y los términos de error de la ecuación en que aparecen están incorrelacionados entre sí:  $E(\xi\delta') = E(\delta\xi') = E(\eta\varepsilon') = E(\varepsilon\eta') = 0$
3. Los términos de error de las variables observadas están incorrelacionados entre sí:  $E(\delta\varepsilon') = E(\varepsilon\delta') = 0$ .

El proceso de estimación consiste en estimar los parámetros de la matriz de varianzas-covarianzas  $\hat{\Sigma}$ , y esta sea lo más parecida posible a la matriz de varianzas-covarianzas muestral  $S$ .

### 2.1.3. Identificación del modelo en los MEC

En el método MEC se dispone de una serie de datos (las varianzas-covarianzas muestrales de las variables observadas) y con ellos se ha de estimar una serie de parámetros (cargas factoriales, varianzas-covarianzas de los factores comunes, y varianzas-covarianzas de los factores específicos o errores). Al igual que ocurre en un sistema de ecuaciones lineales, se puede disponer en principio de más ecuaciones que incógnitas, del mismo número o de mayor número de incógnitas que ecuaciones. La identificación en el modelo MEC hace referencia a si los parámetros del modelo

pueden o no ser determinados de forma única. Si no se imponen restricciones a los parámetros por estimar, habrá un número infinito de soluciones posibles.

Siguiendo a Hatcher (1994) y Ullman (1996), el investigador debería centrarse en las siguientes tareas:

1. Comparar el número de datos con el número de parámetros que han de estimarse. Los datos son siempre las varianzas-covarianzas muestrales, y se ha visto que existen  $q(q+1)/2$ . Como el número de parámetros para estimar es  $[qs + s(s+1)/2 + q(q+1)/2]$ , el modelo estará sin identificar si no se imponen al menos  $qs + s(s+1)/2$  restricciones.
2. Establecer una escala para los factores comunes. Esto se consigue fijando la varianza de cada factor común a uno o el coeficiente de regresión (carga factorial) de una de las variables observadas que cargan sobre cada factor a uno. Si esto no se hace, se produce el denominado problema de indeterminación entre la varianza y las cargas factoriales.
3. Asegurar la identificabilidad de la parte del modelo que contiene la relación entre las variables observadas y los factores. Para esto debe analizarse el número de factores y el número de variables observadas que cargan sobre cada factor o lo componen. Si solo hay un factor, el modelo puede estar identificado si el factor tiene al menos tres variables con cargas no nulas sobre él. Si hay dos o más factores, examínese el número de variables observadas de cada factor. Si cada factor tiene tres o más variables que cargan sobre él, el modelo puede estar identificado si los errores asociados con los indicadores no están correlacionados entre sí, cada variable carga solo sobre un factor y los factores pueden covariar entre ellos. Si solo hay dos indicadores por factor, el modelo puede estar identificado si los errores asociados con cada indicador no están correlacionados, cada indicador carga solo sobre un factor y ninguna de las covarianzas entre los factores es igual a cero.
4. Fijar arbitrariamente el coeficiente de regresión del error al valor 1<sup>2</sup>.

#### 2.1.4. Estimación del modelo en los MEC

Una vez identificado el modelo, este puede estimarse por cualquiera de los procedimientos señalados: máxima verosimilitud (ML), mínimos cuadrados generalizados (GLS), mínimos cuadrados no ponderados (ULS) o mínimos cuadrados ponderados (WLS o ADF). Las estimaciones de los parámetros se realizan de forma que se minimice la diferencia entre la matriz de varianzas-covarianzas muestrales  $S$  y la matriz de varianzas-covarianzas predicha  $\hat{\Sigma}$ . Cada procedimiento de estimación define de una manera distinta esa diferencia o función de ajuste.

Los procedimientos de estimación se describen a continuación.



### Mínimos cuadrados no ponderados

ULS (*Unweighted Least Squares*) toma como estimadores los valores que minimizan la siguiente función de ajuste:

$$F_{ULS}(S; \Sigma^*) = \frac{1}{2} \text{tr}[S - \Sigma^*]^2 \quad (3)$$

Donde  $\text{tr}$  es la traza de la matriz resultante de la operación subsiguiente, es decir, la suma de los elementos de su diagonal,  $S$  es la matriz de varianzas-covarianzas muestral, y  $\Sigma^*$  es la matriz de varianzas covarianzas estimada que si cumple con las restricciones realizadas para la identificación del modelo.

### Mínimos cuadrados generalizados

GLS (*Generalized Least Squares*) se base en minimizar la siguiente función de ajuste:

$$F_{GLS}(S; \Sigma^*) = \frac{1}{2} \text{tr}[(S - \Sigma^*)S^{-1}]^2 \quad (4)$$

### Máxima verosimilitud

ML (*Maximun Likelihood*) implica minimizar la siguiente función de ajuste:

$$F_{ML}(S; \Sigma^*) = \frac{1}{2} \text{tr}[S(\Sigma^*)^{-1}] + \log |\Sigma^*| - \log |S| - q \quad (5)$$

Donde  $q$  es el número de variables observadas.

### Libre distribución asintótica o mínimos cuadrados ponderados

ADF (*Asymptotically Distribution Free*) o WLS (*Weigthed least squares*) minimiza una función definida mediante la siguiente expresión:

$$F_{ADF}(S; \Sigma^*) = (s - \sigma(\Theta))'W^{-1}(s - \sigma(\Theta)) \quad (6)$$

Donde  $s$  es el vector de datos, es decir, la matriz de varianzas-covarianzas muestrales, pero escrita como un solo vector;  $\sigma$  es la matriz de varianzas-covarianzas estimada en forma de vector y donde con el término  $(\Theta)$  se ha querido indicar que se deriva de los parámetros del modelo (coeficientes de regresión, varianzas-covarianzas).  $W$  es una matriz que pondera las diferencias cuadráticas entre las matrices de varianzas-covarianzas muestrales y estimadas.

#### 2.1.5. Bondad de ajuste del modelo estimado

Al igual que ocurría con la identificación del modelo, los principales indicadores de medida de ajuste son los mismos que los usados en el AFC. Es necesario determinar hasta qué punto el modelo asumido se ajusta a los datos muestrales. Si se detectan problemas de ajuste, será necesario reespecificar el modelo. Los criterios se describen a continuación.

**Matriz residual de covarianzas.** La matriz residual de covarianzas es la diferencia entre la matriz de varianzas-covarianzas poblacional estimada y la muestral. Esta debe acercarse lo más posible a la matriz nula y sus valores deben estar homogéneamente distribuidos.

**Estadísticos  $\chi^2$  para el contraste global del modelo.** Se establece la siguiente hipótesis nula:  $H_0 : \Sigma_{nc} = \Sigma$ , donde  $\Sigma_{nc}$  es la matriz de varianzas-covarianzas de  $x$  no condicionada al modelo. La estimación de esta matriz es la matriz muestral  $S$ ,  $\Sigma$  es la matriz de varianzas-covarianzas del vector  $x$  condicionada al modelo. La hipótesis alternativa postula que la matriz  $\Sigma_{nc}$  es igual a cualquier matriz que sea definida positiva. Si el modelo es el adecuado se puede esperar que se rechace la hipótesis nula.

Los supuestos para usar los estadísticos  $\chi^2$  son:

- Se asume la hipótesis de normalidad de las variables observadas.
- El análisis se basa en la matriz de varianzas-covarianzas y no en la de correlaciones.
- El tamaño muestral es lo suficientemente grande para que se justifiquen las propiedades asintóticas del contraste.

Como estos requisitos se cumplen rara vez, Bentler & Bonett (1980), Long (1983) y Ullman (1996) dicen que el estadístico debe aplicarse con precaución con muestras grandes, dado que incluso pequeñas diferencias entre las matrices de covarianza muestral y la estimada serán evaluadas como significativas.

**Estadísticos ad hoc.** Esta limitación ha dado lugar al desarrollo de varios estadísticos ad hoc de bondad de ajuste (ver Tabla 3).

### 2.1.6. Interpretación del modelo y su reespecificación

La interpretación del modelo pasa por establecer qué parámetros estimados han resultado significativos, y luego reespecificar el modelo eliminando los no significativos. Se examina si los estimadores de los parámetros son razonables en dos sentidos (Uriel & Aldás 1996): 1) ¿Toman valores adecuados teóricamente? y 2) ¿Son significativos?

Para que exista compatibilidad de los valores que toman los parámetros estimados con el modelo estadístico, las respuestas a estas preguntas deben ser negativas:

- ¿Existen correlaciones superiores a la unidad?
- ¿Existen cargas factoriales estandarizadas fuera del intervalo -1, 1?
- ¿Los residuos estandarizados son anormalmente grandes o pequeños?
- ¿Hay estimaciones negativas de las varianzas?

Tabla 3: *Índices ad hoc de bondad de ajuste para los MEC. Fuente: elaborados a partir de Uriel & Aldás (2005)*

Grupo	Índice	Sigla	Valor óptimo
Índices comparativos  de ajuste	Normed Fit Index	NFI	> 0.9
	Nonnormed Fit Index	NNFI	> 0.9
	Comparative Fit Index	CFI	> 0.9
	Incremental Fit Index	IFI	> 0.9
	Índice de ajuste absoluto	MFI	> 0.9
Índices de proporción  de varianza	Goodness of Fit Index	GFI	> 0.9
	Adjusted Goodness of Fit Index	AGFI	> 0.9
Índices de grado  de parsimonia	Akaike Information Criterion	AIC	Comp. con modelo independiente
	Consistent AIC	CAIC	Comp. con modelo independiente
Índices basados  en residuos	Root Mean Squared Residual	RMR	< 0.1
	RMR Error of Approximation	RMSEA	< 0.1

Si existen respuestas no negativas, y aunque el ajuste global del modelo fuera óptimo, existiría una incompatibilidad que puede haberse originado por los siguientes motivos (Long 1983):

1. El modelo está mal especificado.
2. Los datos no respaldan la hipótesis de normalidad multivariante de las variables observadas.
3. La muestra es demasiado pequeña.
4. El modelo está demasiado cerca de no estar identificado, lo que dificulta o desestabiliza la estimación de algunos parámetros.
5. Los valores perdidos de algunas variables observadas han provocado que cada elemento de la matriz de covarianzas muestral esté calculado sobre una muestra diferente.

## 2.2. Comparación de los métodos de estimación de los MEC

Ullman (1996) recomienda:

- Los métodos ML y GLS son la mejor opción con pequeñas muestras siempre que sea plausible la asunción de normalidad e independencia (entre los términos de error y los factores comunes).
- En el caso de que ambos supuestos no parezcan razonables, se recomienda recurrir a la estimación ML denominada *escalada*.

Sin embargo, no todos los programas estadísticos presentan en sus opciones al método ML *escalada*. En esta investigación se compararon métodos que están presentes en el programa AMOS (*Analysis of Moment Structures*). Este módulo del SPSS permite especificar, estimar, evaluar, presentar y perfeccionar modelos de actitud y comportamiento que muestran relaciones complejas hipotéticas entre variables, mediante una interfaz gráfica de usuario de forma sencilla, usando el modelado de ecuaciones estructurales (SEM, por sus siglas en inglés).

Un análisis comparativo de los distintos métodos se muestra en la Tabla 4.

## 2.3. Modelos de relación de la calidad de servicio con otros constructos

Brady et al. (2002) señalaron que la calidad de servicio y la satisfacción no explican la totalidad de la variación en las intenciones de compra de los consumidores. Factores como el valor de un servicio, el entorno físico en el que la prestación del servicio se produce y factores incontrolables, como el comportamiento de otros clientes en el lugar del servicio, merecen un examen adicional debido a su potencial influencia sobre la disposición de los consumidores para hacer compras.

Brady et al. (2005) identificaron cuatro modelos competitivos de evaluación del servicio que son usados comúnmente para describir los antecedentes de las intenciones de comportamiento (ver Figura 2).

Los modelos son nombrados de acuerdo con sus especificaciones y a las literaturas en las que se encuentran. El primer modelo, el de *valor*, posiciona el valor como un constructo mediador central (Parasuraman & Grewal 2000, Sirohi et al. 1998, Sweeney et al. 1999). El segundo modelo, el de *calidad de servicio*, posiciona la calidad de servicio como un conductor central de las intenciones de comportamiento (Athanasopoulos 2000, Boulding et al. 1993, Zeithaml et al. 1996). La tercera conceptualización es el modelo de *satisfacción* y es similar a los modelos que posicionan a la satisfacción como el determinante clave de las intenciones de comportamiento (Anderson & Fornell 1994, Fornell et al. 1996, Heskett et al. 1994). El cuarto modelo *comprehensivo* adopta una postura diferente, ya que especifica que la calidad del servicio, el valor del servicio y la satisfacción están directamente relacionados en las intenciones de comportamiento. Es *comprehensivo* en el sentido de que la

Tabla 4: *Ventajas y desventajas de los métodos de estimación de los MEC. Fuente: elaboración propia basada en Long (1983), Ullman (1996), Uriel y Aldás (2005), Brown (2006), Ximénez y García (2005) y García (2011)*

Método	Ventajas	Desventajas
ML	Se pueden hallar todos los índices de bondad de ajuste. Se pueden realizar pruebas de significancia $t$ de los parámetros estimados.	Exige la asunción de normalidad entre los términos de error y los factores comunes. Funciona bien para muestras pequeñas.
GLS	Se pueden hallar todos los índices de bondad de ajuste. Se pueden realizar pruebas de significancia $t$ de los parámetros estimados.	Exige la asunción de normalidad entre los términos de error y los factores comunes. Funciona bien para muestras pequeñas. Resulta inapropiado cuando el modelo aumenta en tamaño y complejidad.
ULS	No es necesaria la suposición de normalidad de las variables observadas. Para corregir el problema de la dependencia de las unidades, se debe tomar como datos la matriz de correlaciones.	No se pueden hallar todos los contrastes estadísticos asociados. No se pueden realizar pruebas de significancia $t$ de los parámetros estimados. Los estimadores dependen de la escala de medida de las variables observadas. Puede ser incorrecto si las diferencias de varianzas entre las variables son arbitrarias. Es un método poco utilizado.
WLS, ADF	No es necesaria la suposición de normalidad de las variables observadas. Permite introducir en los análisis variables ordinales, variables dicotómicas y variables continuas que no se ajusten a criterios de normalidad.	El tamaño muestral debe superar $(q(q+1))/2$ , donde $q$ es el número de variables observadas en el modelo. El valor del estadístico chi-cuadrado que proporciona únicamente será preciso cuando la muestra sea lo suficientemente grande.

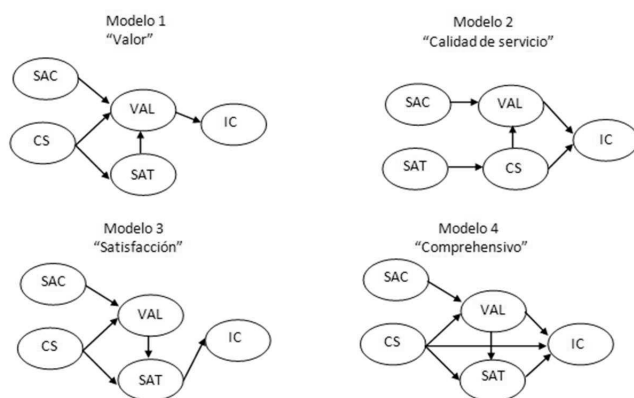


Figura 2: Cuatro modelos de evaluación del servicio -SAC: sacrificio; CS: calidad de servicio; SAT: satisfacción; VAL: valor; e IC: intenciones de comportamiento. Fuente: modificado de Brady et al. (2005)

calidad de servicio, el valor y la satisfacción influyen directa y conjuntamente en las intenciones de comportamiento.

Brady et al. (2005) condujeron dos estudios para probar la validez externa de los cuatro modelos en diversos escenarios. El primer estudio (estudio 1) probó los modelos en varios países (restaurantes de comida rápida y tiendas de abarrotes en cinco países: Australia, Hong Kong, Marruecos, Holanda y los Estados Unidos) y el segundo (estudio 2) solo en Estados Unidos (servicios de aerolíneas y de consulta médica). El examen de los *paths* (relaciones) individuales a lo largo de las muestras sugirió que el modelo comprehensivo es el que tiene mayor validez externa y mejores índices de ajuste.

De la misma manera, Choi et al. (2004) y Chen (2008), entre otros, probaron empíricamente modelos comprehensivos que relacionan la calidad de servicio, el valor y la satisfacción con las intenciones de comportamiento. Los resultados de estos estudios presentaron pruebas de que la secuencia causal del modelo de la actitud multiatributo, es decir, la cognición (la calidad del servicio y valor) - afecta (satisfacción) - al cognado (intención del comportamiento), es robusta. Sin embargo, existieron algunas diferencias. El determinante más importante sobre la satisfacción y las intenciones de comportamiento fue la calidad de servicio en el estudio de Choi et al. (2004), mientras que en el estudio de Chen (2008) fue el valor percibido; además no existió relación entre la calidad de servicio y la satisfacción en el estudio de Chen (2008).

## 2.4. Modelos estructurales de la calidad de servicio de clases

El único estudio que presenta un modelo estructural del servicio de clases publicado es el realizado por Banwet & Datta (2003), desarrollado en un instituto en Nueva

Delhi. La muestra de la encuesta consistió en 168 estudiantes que asistieron a cuatro clases impartidas por un profesor para el mismo curso en un intervalo de dos semanas. El modelo de ecuaciones estructurales (Figura 3) que analizaron muestra el proceso a través del cual varias dimensiones de la calidad percibida de la clase (tangibles, confiabilidad, receptividad, seguridad, empatía y resultado) afectan la satisfacción de los estudiantes y sus intenciones de posclase. Las intenciones de comportamiento de posclase son dependientes de la calidad percibida de clases, así como también de la variable mediadora de la satisfacción del estudiante. El modelo fue probado usando el software AMOS 4.01.

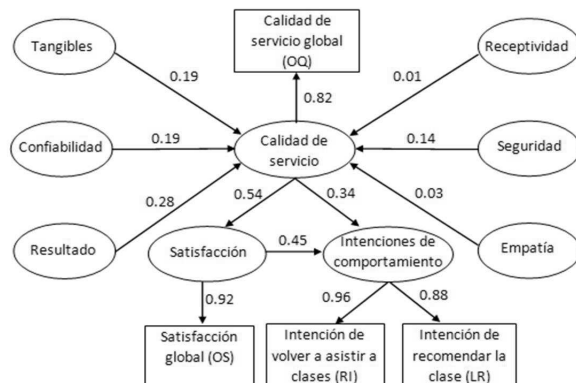


Figura 3: *Modelo estructural de Banwet y Datta (2003). Fuente: modificado de Banwet y Datta (2003)*

En este estudio, se midió la calidad de servicio usando los 22 ítems del instrumento SERVQUAL (desarrollado por Parasuraman et al. (1988)), con una escala de Likert de 7 puntos, que va desde *mucho mejor de lo esperado* hasta *mucho peor de lo esperado*. También incluyeron la dimensión técnica de la calidad de servicio con 4 ítems (resultado del servicio). De esa misma manera midieron la calidad de servicio global (OQ), la satisfacción global de los estudiantes (OS), una valoración de la importancia de cada uno de los 26 ítems de la escala de calidad de la clase, y las intenciones de asistir a otra clase impartida por el mismo profesor (RI) y recomendar a otros a asistir a esas clases (LR).

Los valores de los coeficientes de Cronbach para las sub-escalas iban de 0,72 a 0,87, indicando que la prueba tenía consistencia interna. Los coeficientes de correlación ítem-al-total para los ítems de desempeño individual fueron mayores que el valor 0,35. La fiabilidad ( $\alpha$  de Cronbach) para la combinación lineal de las 5 sub-escalas para las escalas de importancia y desempeño fueron de 0,82 y 0,92, respectivamente. La calidad de la clase fue medida como la brecha entre el desempeño percibido y las expectativas estudiantiles. También probaron la validez predictiva de las escalas y se encontró evidencia favorable. El ajuste global del modelo fue adecuado con el índice de bondad de ajuste ( $GFI = 0,986$ ) y el índice de ajuste normalizado ( $NFI = 0,989$ ).

### 3. Contexto de la investigación

#### 3.1. La Universidad Privada Boliviana (UPB)

La Universidad Privada Boliviana se constituyó en un candidato idóneo para llevar a cabo la presente investigación empírica, ya que es una institución consolidada en el sector de la educación privada boliviana, tiene un sistema de gestión de calidad orientado hacia el alumno y mecanismos de implementación de mejoramiento de la calidad de sus servicios.

#### 3.2. Recopilación, tratamiento y análisis previos de datos

Se realizó un estudio exploratorio por medio del cual se desarrollaron instrumentos específicos para la medición de la calidad de servicio de clases percibida por los alumnos, el valor percibido del servicio de clases, la satisfacción y las intenciones de comportamiento de posclase, con validez de contenido.

Para que esos instrumentos tengan una estructura dimensional adecuada se realizó un análisis factorial exploratorio (AFE) a una muestra de 300 estudiantes y se obtuvieron los siguientes resultados:

- Batería para medir la calidad de servicio global de 4 ítems con una estructura unidimensional, fiabilidad de 0,953 y una varianza total extraída de 88 %.
- Batería para medir el valor percibido de clases de 4 ítems con una estructura unidimensional, fiabilidad de 0,922 ( $\alpha$  de Cronbach) y una varianza total extraída de 81 %.
- Batería para medir la satisfacción del servicio de clases de 4 ítems con una estructura unidimensional, fiabilidad de 0,961 y una varianza total extraída de 89 %.
- Batería para medir las intenciones de comportamiento de posclase de 5 ítems con una estructura unidimensional, fiabilidad de 0,963 y una varianza total extraída de 87 %. El ítem 5 fue eliminado, ya que no fue significativo.
- Por último, en la Tabla 5 se exponen las baterías para la medición de la calidad de servicio global (CSG), el valor percibido del servicio de clases (VAL), la satisfacción estudiantil (SAT) y las intenciones de comportamiento de posclase (IC).

### 4. Idoneidad de los datos

Antes de aplicar los MEC en la investigación, se realizaron algunas pruebas sobre los datos recopilados, en particular, pruebas de idoneidad y conformidad de las muestras al patrón poblacional.



Tabla 5: *Baterías resultantes para la medición de la calidad de servicio global, valor percibido, satisfacción e intenciones de comportamiento de posclase. Fuente: elaboración propia.*

N.º	<b>Calidad de servicio global de clases (CSG)</b>
CSG1	El docente prestó un excelente servicio de clases global
CSG2	El docente impartió un servicio de clases de muy alta calidad
CSG3	El docente otorgó un alto estándar de servicio de clases
CSG4	El docente entregó un servicio de clases superior en todos los sentidos
N.º	<b>Valor percibido de clases (VAL)</b>
VAL1	La enseñanza que recibí del docente ha sido muy valiosa y útil
VAL2	El producto que recibí en esta clase vale más de lo que pagué
VAL3	Valió la pena dedicar mi tiempo a esta clase, por el conocimiento que obtuve
VAL4	Valió la pena gastar mi energía y esfuerzo en esta clase, porque fue una clase de calidad
N.º	<b>Satisfacción global de las clases (SAT)</b>
SAT1	Estoy satisfecho de haber pasado clases con el docente
SAT2	Mi elección de pasar clases con el docente fue sabia
SAT3	Creo que hice lo correcto al cursar esta materia con el docente
SAT4	Creo que mi experiencia con este docente ha sido agradable
N.º	<b>Intenciones de comportamiento de posclase (IC)</b>
IC1	Asistiría a otro curso (materia) impartido por el mismo docente
IC2	Recomendaría y animaría a mis compañeros a tener clases con el docente
IC3	Diré cosas positivas acerca del docente a mis compañeros
IC4	Si fuera una autoridad académica, premiaría al docente por su desempeño
IC5	Permanecería en esta universidad hasta concluir mis estudios, si todos los docentes otorgaran el servicio de calidad que recibí en esta clase

#### 4.1. Pruebas de idoneidad

Se realizaron dos tipos de pruebas de idoneidad: previas a la técnica (confiabilidad de los datos de partida, existencia de valores perdidos y de observaciones) y propias a la técnica (comprobación de normalidad, linealidad, homocedasticidad, independencia de observaciones, etc.)

En el primer caso se eliminaron los cuestionarios mal llenados, inconclusos o que tenían patrones de tendencia anormales (como la aquiescencia -tendencia a estar de acuerdo con el enunciado de un ítem, independientemente de su contenido- y los que tenían doble respuesta en un ítem). Se prepararon 1443 cuestionarios, de los cuales por fallas en la logística se obtuvieron 787 cuestionarios llenados; 187

se eliminaron por presentar deficiencias en el llenado y en las respuestas. De ese modo se obtuvieron 600 cuestionarios aptos para el análisis. Luego se generaron dos muestras, cada una de 300 cuestionarios, tomando en cuenta que en cada una de ellas se encuentren los mismos cursos encuestados, un número similar de estudiantes de sexo masculino y femenino, y un porcentaje similar de estudiantes de las facultades de Ingenierías y Ciencias Empresariales (según el patrón poblacional).

En el segundo caso se comprobaron la normalidad y linealidad, ya que estas dos características son requeridas para el uso de los modelos de estructuras de covarianza. Se realizaron pruebas de normalidad multivariante mediante el método del gráfico chi-cuadrado (Johnson & Wichern 1998), las cuales dieron como resultado que no se cumple el supuesto, tanto para la muestra 1 como para la muestra 2. Este resultado limita el uso de métodos como el ML o el GLS para el análisis de los MEC. Los métodos más adecuados deberían ser el ULS o el WLS-ADF. Por otro lado, se realizó la prueba de linealidad entre los ítems de la calidad de servicio de clases, mediante gráficos de dispersión y se comprobó que las relaciones son aproximadamente lineales.

## 4.2. Conformidad de las muestras al patrón poblacional

Las dos muestras recopiladas para el tratamiento y análisis de datos siguen las tendencias que existen en el patrón poblacional, que está conformado por las características de los estudiantes de la UPB inscritos en el semestre I - 2011 (ver Tabla 6).

Tabla 6: *Conformidad de las muestras al patrón poblacional de la UPB (en porcentaje). Fuente: elaboración propia con base en el sistema de información académico de la UPB.*

Característica	Género		Facultad	
	Masculino	Femenino	Ingenierías	Empresariales
Poblacional	56	44	51	49
Muestra 1	49	51	47	53
Muestra 2	49	51	48	52

Las dos muestras tienen características similares al patrón poblacional y entre sí, y desde ese punto de vista no existe ninguna anomalía para no pensar que son adecuadas para los análisis que se realizaron posteriormente.

## 5. Resultados

Los resultados estandarizados del modelo estructural de evaluación del servicio de clases, tomando en cuenta los distintos métodos de estimación, se muestran en la Figura 4. Se usó el programa AMOS para la generación de los resultados de los MEC.

En la Figura 4 solo se muestran los resultados de las relaciones estructurales y no así los resultados de los componentes de medida de los diferentes constructos. Por otro lado, el modelo comprehensivo fue reespecificado debido a que el *path* del valor con las intenciones de comportamiento (VAL - IC) no fue significativo y fue eliminado del modelo.

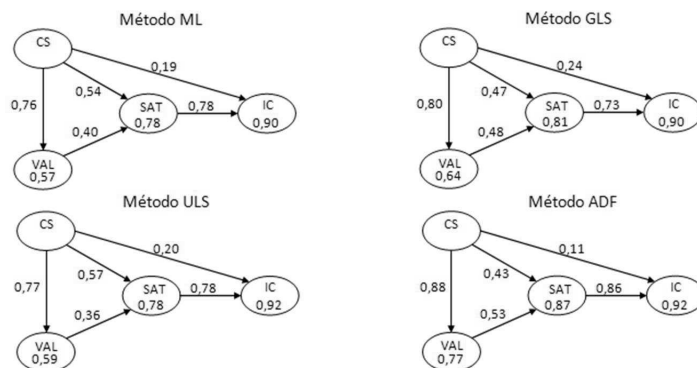


Figura 4: Resultados estandarizados del modelo estructural de evaluación del servicio de clases utilizando varios métodos (CS: calidad de servicio; SAT: satisfacción; VAL: valor; e IC: intenciones de comportamiento). Fuente: Elaboración propia

Si se comparan los valores de los *paths* del modelo estructural de los distintos métodos con los del método ULS (que debería ser el más adecuado), se puede observar lo que está en la Tabla 7.

Tabla 7: Comparación de los valores de los *paths* del modelo de evaluación del servicio de clases con distintos métodos para los MEC, tomando como base el método ULS (en porcentaje). Fuente: elaboración propia.

Path	Métodos			
	ULS	ML	GLS	ADF
CS - IC	0	-5	20	-45
CS - VAL	0	-1	4	14
CS - SAT	0	-5	-18	-25
VAL - SAT	0	11	33	47
SAT - IC	0	0	-6	10

Con el método ADF, la calidad de servicio tiene un efecto débil sobre las intenciones de comportamiento y el menor efecto con la satisfacción, comparativamente al de los otros métodos. Sin embargo, el efecto de la calidad de servicio sobre el valor, del valor sobre la satisfacción y de la satisfacción sobre las intenciones de comportamiento son los mayores al compararlos con los otros métodos.

Con el método GLS se potencia el efecto de la calidad de servicio con las intenciones de comportamiento y se disminuye el efecto de la satisfacción.

Los resultados con el método ML son similares a los del método ULS, potenciando el efecto del valor percibido sobre la satisfacción. En general se puede observar que existen desviaciones porcentuales que van desde -45 % a 47 % de los valores de los *paths* de los distintos métodos comparados con el método ULS. Por lo tanto, la elección correcta del método de estimación de parámetros es muy importante para la interpretación adecuada del modelo MEC que se está considerando.

En la Tabla 8 se muestran las correlaciones múltiples cuadradas de las intenciones de comportamiento de los cuatro métodos.

Tabla 8: *Comparación de los valores de los paths del modelo de evaluación del servicio de clases con distintos métodos para los MEC, tomando como base el método ULS (en porcentaje). Fuente: elaboración propia.*

Constructos	Métodos			
	ML	GLS	ULS	ADF
CS	0.00	0.00	0.00	0.00
VAL	0.57	0.64	0.59	0.77
SAT	0.78	0.81	0.78	0.87
IC	0.90	0.90	0.92	0.92

El mayor valor de R2 corresponde a las intenciones de comportamiento sin importar el método de estimación de parámetros elegido; así que se estima que los predictores de las intenciones de comportamiento (calidad de servicio global, valor percibido y satisfacción) explican el 90 o 92 % de su varianza, dejando que un término de perturbación explique el restante 10 u 8 % de su varianza. Esto indica que el modelo tiene un gran poder predictivo.

La calidad de servicio es una variable predictora, por eso presenta un R2 de cero. Con los métodos GLS y ADF se incrementa el poder predictivo del valor y la satisfacción, más que en los métodos ML y ULS. También se puede observar que con los cuatro métodos se logra un buen poder predictivo de las intenciones de comportamiento con base en sus predictores. Sin embargo, los métodos ULS y ADF logran una correlación múltiple cuadrada máxima de 0.92.

La Tabla 9 muestra los índices de bondad de ajuste del modelo por los 4 métodos. Se observa que cuando se usan los métodos GLS y ADF se obtienen índices de ajuste inadecuados, ya que la mayoría no supera el límite de 0,9. Si bien no existe mucha diferencia en los valores de los *paths* entre los métodos ML y ULS (del -5 al 11 %), se verifica que existe gran diferencia en los índices de ajuste (del 3 al 17 %).

Como los índices de ajuste del modelo son muy diferentes usando los distintos métodos de estimación de parámetros, tanto el supuesto de normalidad multivariante y el tamaño muestral juegan un papel muy importante en la elección del método adecuado. Como se probó que los datos no siguen una distribución normal multivariante, los métodos ML y GLS quedan excluidos. Entonces los métodos ULS y ADF son los más adecuados. Sin embargo, el método ADF exige un número considerable de casos (tamaño muestral de 500 o más) (Uriel & Aldás 1996).

Tabla 9: *Índices ad hoc de bondad de ajuste de los modelos comparados. Fuente: elaboración propia.*

Índices de ajuste		Métodos			
Grupo	Sigla	ML	GLS	ULS	ADF
Chi-cuadrado	$\chi^2$ (114gl)	395.57	305.36	142.57	267.46
Índices comparativos de ajuste	NFI	0.940	0.526	0.998	0.667
	TLI	0.948	0.551	-	0.726
	CFI	0.957	0.624	-	0.770
	IFI	0.957	0.639	-	0.777
Índices de proporción de varianza	GFI	0.867	0.880	0.999	0.873
	AGFI	0.821	0.839	0.998	0.830
Índices de grado de parsimonia	AIC	473.57	383.36	-	345.46
	CAIC	657.027	566.815	-	528.90
Índices basados en residuos	RMR	0.074	0.153	0.056	0.171
	RMSEA	0.091	0.075	-	0.067

El mejor método para el análisis del modelo de evaluación de la calidad de servicio de clases es el ULS, ya que es el único que logra índices de bondad de ajuste adecuados y un gran poder predictivo para el modelo considerado.

## 6. Conclusiones

La elección del método de estimación de parámetros en los MEC es crucial, ya que afecta los valores de los *paths* del modelo que se está investigando, disminuyendo o aumentando las fuerzas de las relaciones entre constructos; el ajuste del modelo a los datos muestrales y el poder predictivo del modelo con base en sus predictores.

Cuando el supuesto de normalidad multivariante no se cumple para los datos muestrales, los métodos ML y GLS resultan inadecuados, ya que el modelo presenta índices de bondad de ajuste inadecuados. El método ADF, si bien tiene la ventaja de poder ser usado sin la suposición de normalidad multivariante, tiene la desventaja de necesitar un tamaño muestral muy grande; en esta investigación se trabajó con un tamaño muestral de 300, que generó índices de bondad de ajuste inadecuados. El método ULS se constituye en la mejor alternativa cuando no se cumple el supuesto de normalidad multivariante, ya que no requiere de tamaños grandes, y se obtiene un buen ajuste del modelo.

Analizando el modelo específico de la investigación de evaluación del servicio de clases, fallar en la elección del método trae serias consecuencias en la implementación de programas de mejoramiento de este servicio, ya que podrían enfocarse en acciones que no tienen el efecto debido para mejorar las intenciones de posclase de los estudiantes.

Usando el método ULS se llega a la conclusión de que la satisfacción es un media-

dor poderoso entre la calidad de servicio de clase y el valor sobre las intenciones de posclase, ya que es mayor el efecto indirecto de la calidad de servicio de clases ( $0,57 \times 0,78 = 0,44$ ) que el efecto directo (0,20) sobre las intenciones de comportamiento de posclase. Lo mismo ocurre con el efecto indirecto del valor percibido ( $0,36 \times 0,78 = 0,28$ ) comparado al efecto directo (no significativo). Eso significa que un plan de gestión para el mejoramiento del servicio de clases, deberá centrarse principalmente en mejorar la satisfacción estudiantil del servicio de clases, más bien que en el mejoramiento de la calidad del servicio de clases y en el mejoramiento de la percepción del valor del servicio de clases.

**Recibido: 16 de noviembre de 2012**

**Aceptado: 28 de febrero de 2013**

## Referencias

- Anderson, E. W. & Fornell, C. (1994), *Service Quality: New Directions in Theory and Practice. Chapter 11: A Customer Satisfaction Research Prospectus*, SAGE Publications, Inc.
- Athanassopoulos, A. D. (2000), 'Customer satisfaction: cues to support market segmentation and explain switching behavior', *Journal of Business Research* **47**, 191–207.
- Banwet, D. K. & Datta, B. (2003), 'A study of the effect of perceived lecture quality on post-lecture intentions', *Work Study* **52**(5).
- Bentler, P. M. & Bonett, D. G. (1980), 'Significance test and goodness of fit in the analysis of covariance structures', *Psychological Bulletin* **88**, 588–606.
- Bollen, K. (1989), *Structural equations with latent variables*, New York: Wiley & Sons.
- Boulding, W., Kalra, A., Staelin, R. & Zeithaml, V. A. (1993), 'A dynamic process model of service quality: from expectations to behavioral intentions', *Journal of Marketing Research* **30**, 7–27.
- Brady, M. K., Cronin, J. J. & Brand, R. R. (2002), 'Performance-only measurement of service quality: a replication and extension', *Journal of Business Research* **55**, 17–31.
- Brady, M. K., Knight, G. A., Cronin, J. J., G., T., Hult, M. & Keillor, B. D. (2005), 'Removing the contextual lens: a multinational, multi-setting comparison of service evaluation models', *Journal of Retailing* **81**(3), 215–230.
- Brown, T. A. (2006), *Confirmatory Factor Analysis for Applied Research*, New York: The Guilford Press.

- Chen, C. F. (2008), 'Investigating structural relationships between service quality, perceived value, satisfaction, and behavioral intentions for air passengers: evidence from Taiwan', *Transportation Research, Part A* **42**, 709–717.
- Choi, K. S., Cho, W. H., Lee, S., Lee, H. & Kim, C. (2004), 'The relationships among quality, value, satisfaction and behavioral intention in health care provider choice: A south korean study', *Journal of Business Research* **57**, 913–921.
- Fornell, C., Johnson, M. D., Anderson, E. W., Cha, J. & Bryant, B. E. (1996), 'The american customer satisfaction index: nature, purpose, and findings', *Journal of Marketing* **60**, 7–18.
- García, M. A. (2011), Análisis causal con ecuaciones estructurales de la satisfacción ciudadana con los servicios municipales, Proyecto fin de máster en técnicas estadísticas, Universidad de Santiago de Compostela, Facultad de Matemáticas.
- Hatcher, L. (1994), *A step by step approach to using the SAS system for factor analysis and Structural Equation Modelling*, SAS Institute.
- Heskett, J. L., Jones, T. O., Loveman, G. W., Sasser, E. W. & Schlesinger, L. A. (1994), 'Putting the service-profit chain to work', *Harvard Business Review* **72**(2), 164–174.
- Johnson, R. A. & Wichern, D. (1998), *Applied multivariate statistical analysis*, Upper Saddle River, NJ: Prentice Hall.
- Jöreskog, K. G. & Sörbom, D. (1989), *Lisrel 7: A guide to the Program Applications*, Chicago: SPSS Inc.
- Long, J. S. (1983), *Covariance structure models: an introduction to LISREL (Quantitative Applications in the Social Sciences)*, SAGE Publications, Inc.
- Parasuraman, A. & Grewal, D. (2000), 'Serving customers and consumers effectively in the twenty-first century: A conceptual framework and overview', *Journal of the Academy of Marketing Science* **28**(1), 9–16.
- Parasuraman, A., Zeithaml, V. A. & Berry, L. L. (1988), 'Servqual: a multiple-item scale for measuring consumer perceptions of service quality', *Journal of Retailing* **64**(1), 12–40.
- Sirohi, N., McLaughlin, E. W. & Wittink, D. R. (1998), 'A model of consumer perceptions and store loyalty intentions for a supermarket retailery', *Journal of Retailing* **74**(2), 223–245.
- Sweeney, J. C., Soutar, G. N. & Johnson, L. W. (1999), 'The role of perceived risk in the quality-value relationship: a study in a retail environment', *Journal of Retailing* **75**(1), 77–105.

- Ullman, J. B. (1996), *Structural equation modeling. In B. G. Tabachnick & L. S. Fidell, Using Multivariate Statistics*, New York: Harper Collins.
- Uriel, E. & Aldás, J. (1996), *Análisis multivariante aplicado*, Thomson (ed.), Madrid, España.
- Ximénez, M. C. & García, A. G. (2005), ‘Comparación de los métodos de estimación de máxima verosimilitud y mínimos cuadrados no ponderados en el análisis factorial confirmatorio mediante simulación monte carlo’, *Psicothema* **17**(3), 528–535.
- Zeithaml, V. A., Berry, L. L. & Parasuraman, A. (1996), ‘The behavioral consequences of service quality’, *Journal of Marketing* **60**(2), 31–46.