

UNIVERSIDAD NACIONAL AGRARIA LA MOLINA

FACULTAD DE ECONOMIA Y PLANIFICACIÓN

DEPARTAMENTO ACADÉMICO DE ESTADÍSTICA E INFORMÁTICA



TRABAJO MONOGRAFICO

**DESCRIPCIÓN METODOLÓGICA DEL MODELO DE ECUACIONES
ESTRUCTURALES CON EL MÉTODO DE ESTIMACIÓN DE MINIMOS
CUADRADOS PARCIALES**

Presentado para optar el título de Ingeniero Estadístico e Informático

EDITH RITA MAMANI TONE

Modalidad Examen Profesional

Lima – Perú

2017

Dedicatoria

A mi Padre Dios por su protección y provisión de cada día.

A mi Esposo Alberto, Padres y familia por su paciencia y apoyo incondicional.

Agradecimientos

A las autoridades de la Universidad Agraria La Molina.

A mi asesor por su constancia y tiempo dedicado a la revisión del presente trabajo.

A los profesores del Dpto. de Estadística e Informática por sus conocimientos impartidos, calidad humana y profesionalismo.

INDICE

RESUMEN

INTRODUCCION	1
I. MODELO DE ECUACIONES ESTRUCTURALES (SEM)	3
1.1. Ventajas de los modelos de ecuaciones estructurales:	3
1.2. Elementos de un modelo de ecuaciones estructurales	4
1.3. Formulación del modelo	5
1.4. Tipos de variables en los modelos de ecuaciones estructurales	6
1.5. Diagrama de Ruta.....	7
1.6. Características de los modelos de ecuaciones estructurales	10
1.7. Fases o pasos de un modelo de ecuaciones estructurales	11
1.8. Técnicas estadísticas para el análisis de Modelos de Ecuaciones estructurales SEM.....	14
1.8.1. SEM basados en covarianzas	14
1.8.2. SEM basados en componentes.	15
1.8.3. Comparativo entre las principales técnicas estadísticas de SEM	15
II. LOS MODELOS DE ECUACIONES ESTRUCTURALES APLICANDO LA TÉCNICA DE MÍNIMOS CUADRADOS PARCIALES PLS	17
2.1. Principales características del PLS.....	17
2.2. Principales ventajas de PLS	18
III. METODOLOGÍA.....	20
IV. CASO EXPLICATIVO.....	25
4.1. Ejemplo.....	25
V. CONCLUSIONES	31
VI. REFERENCIAS BIBLIOGRÁFICAS	32
ANEXO.....	36

ÍNDICE DE TABLAS

Tabla 1 Diferencia entre PLS y CBM.....	15
Tabla 2: Cargas y cargas cruzadas de las variables latentes	27
Tabla 3 Coeficientes del modelo de medida	28
Tabla 4 Análisis de validez discriminante.....	28
Tabla 5 Resultados del modelo estructural	29
Tabla 6 Índices de ajuste del modelo	30
Tabla 7 Resumen de hallazgos	30

ÍNDICE DE FIGURAS

Figura 1: Modelo de medida y estructural	5
Figura 2: Convenciones Diagrama de rutas	8
Figura 3: Representación de una regresión simple	8
Figura 4: Representación de una regresión múltiple sin relaciones entre los predictores (izquierda) o con relaciones entre los predictores (derecha).	9
Figura 5: Representación gráfica de un modelo de análisis de ruta	9
Figura 6: Representación gráfica de modelo de ruta con una mediadora y mediación total	10
Figura 7: Representación gráfica de modelo de ruta con una mediadora y mediación parcial	10
Figura 8: Metodología Propuesta.....	20
Figura 9: Modelo propuesto e hipótesis	26
Figura 10: Resultados del modelo de ecuaciones estructurales.....	29

RESUMEN

La presente monografía estudia la Descripción Metodológica del Modelo de Ecuaciones Estructurales con el Método de Estimación de Mínimos Cuadrados Parciales.

En el primer capítulo se describe detalladamente la sustentación teórica del Modelo de Ecuaciones Estructurales; en el capítulo 2 se describe el Método de Mínimos Cuadrados Parciales (PLS); en capítulo 3 se describe la Metodología detallando los pasos del Modelo de Ecuaciones Estructurales por el Método de Mínimos Cuadrados Parciales y en el capítulo 4 se observa la Aplicación, paso a paso en un caso para ejemplificar la metodología.

Se concluyó los modelos de ecuaciones estructurales SEM es una extensión de la regresión múltiple, se aplica esta técnica para encontrar relaciones entre variables observables y no observables llamadas (latentes) para pasar posteriormente a estimar los parámetros. tiene como objetivo la predicción, no es preciso que los datos provengan de una distribución normal y puede aplicarse a estudios de muestras pequeñas, permite estimar modelos muy complejos con muchas variables latentes y medibles.

En la aplicación de la descripción metodológica se realizó un estudio empírico el segundo semestre de 2013 sobre una muestra correspondiente a 300 alumnos universitarios chilenos con acceso a bases de datos científicas. Para los cálculos de PLS se utilizó el software WarpPLS 4.0.

Los resultados del ejemplo en el análisis de PLS del caso indicaron la buena capacidad predictiva del modelo de investigación, y a su vez, la explicación del análisis logró ejemplificar en forma clara la metodología propuesta.

Palabras Clave: Modelos de Ecuaciones Estructurales, Método de mínimos cuadrados parciales.

SUMMARY

The present monograph studies the Methodological Description of the Model of Structural Equations with the Method of Estimation of Partial Least Squares.

In the first chapter the theoretical support of the Structural Equation Model is described in detail; Chapter 2 describes the Partial Least Squares Method (PLS); In Chapter 3 the Methodology is described, detailing the steps of the Structural Equation Model by the Partial Least Squares Method and in Chapter 4 the Application is observed, step by step in a case to exemplify the methodology.

The structural equation models were concluded. SEM is an extension of the multiple regression, this technique is applied to find relations between observable and unobservable variables called (latent) to later pass to estimate the parameters. It has as its objective the prediction, it is not necessary that the data come from a normal distribution and it can be applied to studies of small samples, it allows to estimate very complex models with many latent and measurable variables.

In the application of the methodological description, an empirical study was conducted in the second semester of 2013 on a sample corresponding to 300 Chilean university students with access to scientific databases. The WarpPLS 4.0 software was used for the PLS calculations.

The results of the example in the PLS analysis of the case indicated the good predictive capacity of the research model, and in turn, the explanation of the analysis was able to clearly exemplify the proposed methodology.

Keywords: Structural Equation Models, Partial Least Squares Method.

INTRODUCCIÒN

El modelo de ecuaciones estructurales (SEM) es el modelado simultaneo de relaciones entre múltiples constructos (variables latentes) independientes y dependientes.

(Fornell 1982: 289-324) denomina al SEM como análisis multivariantes de segunda generación. La característica común de las técnicas que se acogen bajo esta etiqueta es el reconocimiento metodológico de que la teoría científica implica tanto variables empíricas como abstractas. En este sentido, el propósito de los análisis multivariantes de segunda generación es ayudar a vincular datos y teoría.

Se tiene dos métodos de estimación de los SEM, el método de covarianzas (CBM) y el método de los mínimos cuadrados parciales (PLS). El CBM son modelos basados en la covarianza, adecuados para confirmación de teorías (análisis confirmatorios). PLS son modelos basados en la varianza, los cuales se adaptan mejor para aplicaciones predictivas de desarrollo de teorías (análisis exploratorio) y en la confirmación de teorías (análisis confirmatorio). El principal problema que se encuentra por el método de covarianzas es que no se aplica correctamente para estudios con muestras pequeñas que habitualmente no cumple la distribución normal multivariada y con la independencia de las observaciones. Es así que el método de mínimos cuadrados parciales se adapta bien al análisis de muestras de datos pequeños y datos que no muestran necesariamente la distribución normal multivariante.

Los mínimos cuadrados parciales PLS, es una técnica que busca maximizar la varianza explicada de las variables latentes o constructos. El PLS valora un modelo causal que involucra múltiples variables con múltiples ítems observados, esta valoración se realiza a la vez sobre el modelo estructural (causalidad entre independientes y dependientes constructos) y sobre el modelo de medida (carga de los ítems observados con sus respectivos constructos).

Entre las características destacables de PLS están que no requiere necesariamente una fuerte base teórica (soporta tanto investigación exploratoria como confirmatoria), y es relativamente robusta a desviaciones de normalidad. El objetivo de la presente investigación es hacer una Descripción Metodológica del Estudio de los Modelos de Ecuaciones Estructurales con el Método de Estimación de los Mínimos Cuadrados Parciales.

La presente monografía está conformado por el capítulo 1 donde se describe la sustentación teórica del modelo de ecuaciones estructurales; en el capítulo 2 se describe el método de mínimos cuadrados parciales (PLS); en capítulo 3 se describe la metodología detallando los pasos del modelo de ecuaciones estructurales por el método de mínimos cuadrados parciales y en el capítulo 4 se observa la aplicación, paso a paso en un caso para ejemplificar la metodología, el paso 1 descripción del modelo, el paso 2 validez y fiabilidad del modelo de medida, paso 3 valoración del modelo estructural.

I. MODELO DE ECUACIONES ESTRUCTURALES (SEM)

Definición:

No existe una definición única del concepto de modelo de ecuaciones estructurales, existe numerosas definiciones que se han ido formulando a lo largo del tiempo de la que podemos destacar la de Byrne (1998) “un modelo de SEM es una metodología estadística que utiliza un enfoque confirmatorio del análisis multivalente aplicado a una teoría estructural relacionada con un fenómeno determinado”. La técnica de modelado de ecuaciones estructurales se considera una extensión de varias técnicas multivariantes como la regresión múltiple, el análisis factorial principalmente y el análisis de senderos. (Fernández 2004:218). (Fornell 1982: 289-324) denomina a estos modelos de ecuaciones estructurales como análisis multivariantes de segunda generación. La característica común de las técnicas que se acogen bajo esta etiqueta es el reconocimiento metodológico de que la teoría científica implica tanto variables empíricas como abstractas. En este sentido, el propósito de los análisis multivariantes de segunda generación es ayudar a vincular datos y teoría.

El SEM incluye metodologías estadísticas utilizadas para estimar cadenas de relaciones causales, definidas en modelos teóricos, relacionando dos o más variables latentes, cada una medida a través de un número de variables manifiestas o indicadores (Esposito et al. 2010: 47).

1.1. Ventajas de los modelos de ecuaciones estructurales:

Ruiz y San Martín (2010) señalan la gran ventaja de este tipo de modelos que permiten proponer el tipo y la dirección de las relaciones que se espera encontrar entre las variables contenidas en el, para pasar posteriormente a estimar los parámetros. Por esta razón se denomina modelos confirmatorios.

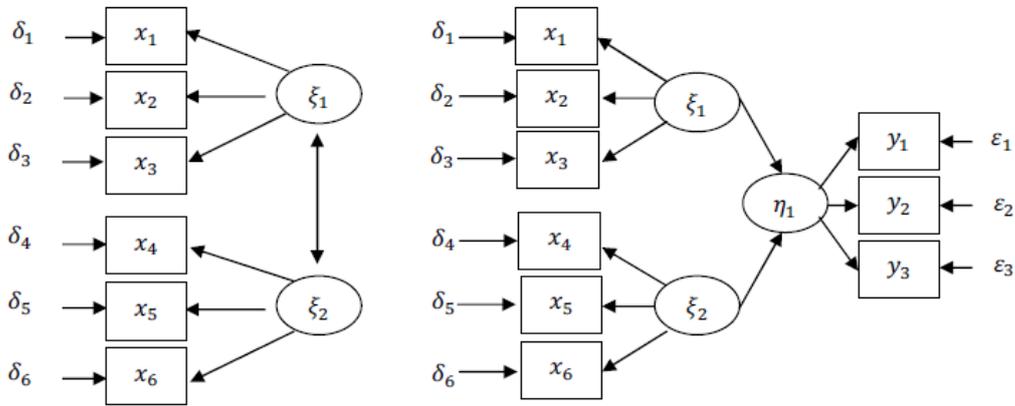
- Permite evaluar modelos teóricos, siendo una de las herramientas más potentes para el estudio de las relaciones causales sobre datos no experimentales cuando estas relaciones son de tipo lineal.
- Permite modelado simultáneo de relaciones entre múltiples constructos dependientes e independientes (que pasan a denominarse variables latentes exógenas y endógenas).
- Permite al investigador construir variables no observables medidas por indicadores, así como estimar el error de las variables observadas.
- Ayuda a seleccionar hipótesis causales relevantes, eliminando aquellas no sustentadas por la evidencia empírica.

1.2. Elementos de un modelo de ecuaciones estructurales

Un modelo de ecuaciones estructurales completo consta de dos partes fundamentales: el modelo de medida y el modelo de relaciones estructurales. El modelo de medida contiene la manera en que cada constructo, o variable latente, está medido mediante sus indicadores observables, los errores que afectan a las mediciones.

Por su parte, el modelo de relaciones estructurales es el que realmente se desea estimar. Contiene los efectos y relaciones entre los constructos. Es similar a un modelo de regresión, pero puede contener además efectos concatenados y bucles entre variables. Además, contiene los errores de predicción (que son distintos de los errores de medición). (García 2011:14).

Figura 1: Modelo de medida y estructural



Fuente: García (2011)

1.3. Formulación del modelo

El modelo de ecuaciones estructurales está conformado por el modelo de medida y el modelo estructural.

Para poder obtener una medida de los constructos o variables latentes, es necesario plantear las expresiones que relacionan cada uno de estos conceptos con sus indicadores, es decir, elaborar el sub modelo estructural y el sub modelo de medida.

1.3.1. Modelo de medición: Especifica las ecuaciones que vinculan las variables latentes a las observadas o indicadores x e y , expresado de forma matricial sería:

$$x = \Lambda_x \xi + \delta$$

$$y = \Lambda_y \eta + \varepsilon$$

1.3.2 Modelo de estructural: Especifica las ecuaciones causales lineales entre las variables latentes del modelo

$$\eta = \beta\eta + \tau\xi + v$$

Siendo:

η : variables endógenas

β : matriz de coeficientes de las variables endógenas.

τ : matriz de coeficientes de la variable exógena

ξ : variable exógena

v : termino de perturbación aleatoria

1.4. Tipos de variables en los modelos de ecuaciones estructurales

Tomás y Fernández (2016). Definen a los modelos de ecuaciones estructurales como un conjunto de técnicas estadísticas multivariantes del modelo lineal general para relacionar de forma compleja una o más variables independientes (o predictoras) con una o más variables dependientes (o criterios). Las variables independientes pueden ser continuas o discretas –como en regresión- y se permite utilizar tanto variables observables como factores. Los factores se estiman a partir de indicadores observables. Hay modelos de ecuaciones estructurales con y sin factores (también conocidos como variables latentes).

En un modelo de ecuaciones estructurales se distinguen distintos tipos de variables según sea su papel y según sea su medición (García 2011:15-16).

- Variable observada o indicador: Variable que se mide a los sujetos. Por ejemplo, las preguntas de un cuestionario.
- Variable latente: Característica que se desearía medir pero que no se puede observar. En consecuencia, que está libre de error de medición. Por ejemplo, una dimensión de un cuestionario o un factor en un análisis factorial exploratorio.
- Variable error: Representa tanto los errores asociados a la medición de una variable como el conjunto de variables que no han sido contempladas en el

modelo y que pueden afectar a la medición de una variable observada. Se considera que son variables de tipo latente por no ser observables directamente.

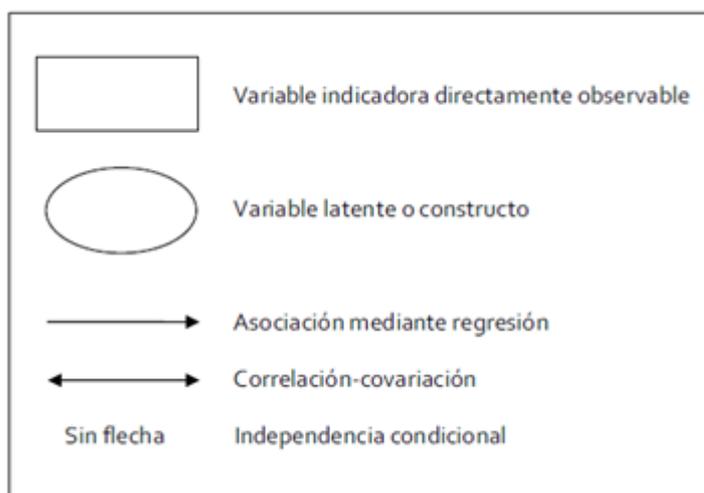
- El error asociado a la variable dependiente: Representa el error de predicción.
- Variable de agrupación: Variable categórica que representa la pertenencia a las distintas subpoblaciones que se desea comparar. Cada código representa una subpoblación.
- Variable exógena: Variable que afecta a otra variable y que no recibe efecto de ninguna variable. Las variables independientes de un modelo de regresión son exógenas.
- Variable endógena: Variable que recibe efecto de otra variable. La variable dependiente de un modelo de regresión es endógena. Toda variable endógena debe ir acompañada de un error.

1.5. Diagrama de Ruta

Para representar un modelo causal y las relaciones que se desea incluir en él se acostumbra a utilizar diagramas similares a los diagramas de flujo. Un diagrama de ruta es una serie de símbolos que permiten poner a prueba “casi” cualquier teoría con tal de que se hayan medido las variables.

- Un cuadrado indica una variable observable.
- Un círculo (elipse) indica un factor o variable latente.
- Una flecha indica un efecto hipotetizado en una dirección (p. ej. de la variable A a la B)
- Una flecha de doble dirección indica una correlación (las variables se relacionan, pero no predice una a la otra).

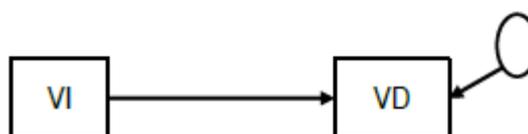
Figura 2: Convenciones Diagrama de rutas



Fuente: García (2011).

Algunos ejemplos de diagramas de ruta, incrementándose en dificultad: El primero simplemente sería el equivalente a una regresión simple (Figura 3) en que una variable independiente (predictor) se supone que afecta (o predice) a una variable dependiente (o criterio). El círculo afectando a la variable dependiente se incluye usualmente en los modelos estructurales para reconocer que no siempre (de hecho casi nunca) es posible predecir o explicar completamente la variable dependiente y queda, por tanto, varianza por explicar o varianza de error (otras causas o azar), pero debido a otros factores no observables, de ahí que se emplee la elipse. A menudo los errores se eliminan del diagrama por simplicidad. En este caso tanto la VI como la VD son variables observables. (Tomás y Fernández 2016:50-51).

Figura 3: Representación de una regresión simple



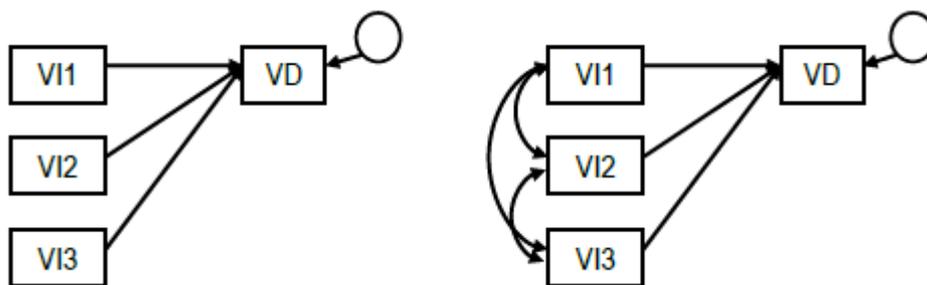
Fuente: Tomás y Fernández (2016)

El caso de una regresión múltiple (otro caso simple de modelo de ecuación estructural), como el que se presenta a continuación. En estos dos diagramas de ruta (Figura 4) se pueden ver que tres VI (VI1-VI3) explican a una VD. En el caso del diagrama de ruta de la

izquierda las tres variables independientes predicen a la variable dependiente, como puede verse por la direccionalidad de las flechas. En el caso del diagrama de ruta de la derecha, también se trata de la predicción de la variable dependiente por las mismas tres variables independientes, pero en este caso

Se ha añadido tres flechas bidireccionales, que indican que, en la teoría mantenida por el investigador, los tres predictores se piensa que pueden estar correlacionados.

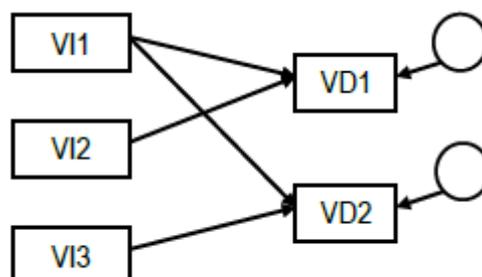
Figura 4: Representación de una regresión múltiple sin relaciones entre los predictores (izquierda) o con relaciones entre los predictores (derecha).



Fuente: Tomás y Fernández (2016)

Los modelos de ecuaciones estructurales pueden mantener un número considerable de variables dependientes y variables dependientes a distintos niveles. Ejemplo de estos tipos de diagrama de ruta a continuación, sería el siguiente con dos variables dependientes (Figura 5).

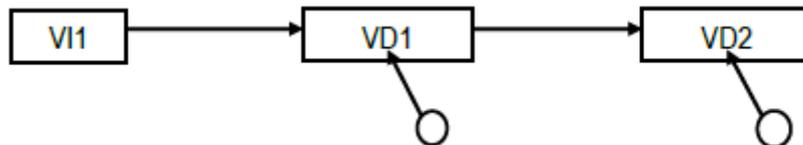
Figura 5: Representación gráfica de un modelo de análisis de ruta



Fuente: Tomás y Fernández (2016)

De la misma forma se puede plantear procesos de mediación: Una variable afecta a otra pero solamente a través de sus efectos sobre una mediadora, lo que se conoce como mediación total, como es el caso que se presenta en el siguiente diagrama de ruta (Figura 6).

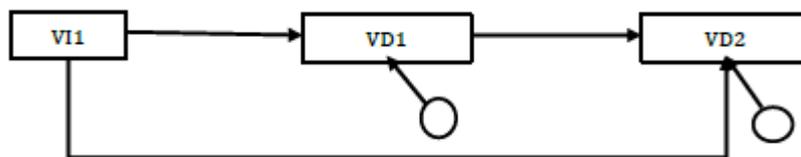
Figura 6: Representación gráfica de modelo de ruta con una mediadora y mediación total



Fuente: Tomás y Fernández (2016)

O bien, se puede plantear que la mediación no es total, sino parcial. Esto es, que hay tanto un efecto directo como uno indirecto de la variable independiente (Figura 7).

Figura 7: Representación gráfica de modelo de ruta con una mediadora y mediación parcial



Fuente: Tomás y Fernández (2016)

1.6. Características de los modelos de ecuaciones estructurales

Según Batista y Coenders (2000), los SEM se caracterizan por dos elementos principales. El primero, evaluar las relaciones de dependencia tanto múltiple como cruzadas. El segundo, el grado para representar conceptos no observados en estas relaciones y tener en cuenta el error de medida en el proceso de estimación.

Para Chin (1998), el sistema de ecuaciones estructurales tiene la ventaja, sobre otros sistemas y técnicas multivariantes, el analizar las relaciones por cada subconjunto de variables, permitiendo también una interrelación entre variables de diferentes grupos,

dependiendo del propósito de la investigación.

Los SEM trabajan con variables observables o medibles (aquellas que tienen un valor de entrada) y una o varias variables latentes o no observadas (que no tiene valor como tal y que puede utilizarse como un concepto), fortaleciendo las correlaciones utilizadas y realizando estimaciones más precisas de los coeficientes estructurales.

1.7. Fases o pasos de un modelo de ecuaciones estructurales

Cupani (2012) y Casas (2014) autores de SEM coinciden en que son seis las fases para aplicar esta técnica. La especificación, identificación, estimación de parámetros, evaluación del ajuste, reespecificación del modelo y la interpretación de resultados que lo conforman.

A continuación, se detalla cada fase:

- a. La especificación es la fase en donde el investigador establece la relación hipotética entre las variables latentes y las observadas, la misma que con el análisis se obtendrán las relaciones correctas.

En esta fase el investigador aplica sus conocimientos teóricos del fenómeno estudiado al planteamiento de las ecuaciones matemáticas relativas a los efectos causales de las variables latentes y a las expresiones que las relacionan con los indicadores o variables observables. Además, se formulan enunciados sobre el conjunto de parámetros, decidiendo entre los que serán *libres* para ser estimados o fijos, a los que se les asignará un valor dado, normalmente cero. Asimismo, en esta etapa se especifican los supuestos estadísticos sobre las fuentes de variación y en concreto sobre la forma de distribución conjunta, que en la mayoría de las técnicas empleadas se considera normalidad multivariante. Por último, se precisará el comportamiento de las variables no incluidas en el modelo, cuyo efecto se recoge en los términos del error de medida o de perturbación.

La claridad del modelo viene determinada por el grado de conocimiento teórico que posea el investigador sobre el tema de estudio, si la información es poco exhaustiva o detallada, la asignación de los parámetros será confusa a priori, por lo que el investigador debe realizar diversos análisis exploratorios de los datos hasta configurar

el modelo, y efectuar el análisis confirmatorio del mismo. (Casas 2002:3)

- b. En la fase de identificación se estiman los parámetros del modelo. Se determina si un modelo está identificado mediante una expresión algebraica que lo demuestre, en función de las varianzas y covarianzas muestrales.

Si el modelo teórico creado es correcto, se procede a la identificación del modelo, en donde nos debemos asegurar que pueden ser estimados los parámetros del modelo. *El modelo está identificado si todos los parámetros lo están.* (Batista, J. y Coenders, G. 2000:68) es decir, si existe una solución única para cada uno de los parámetros estimados.

Determinar si un modelo está identificado debe estudiarse antes de la recogida de datos, en donde se debe comprobar que al menos se dispone para cada parámetro de una expresión algebraica que lo exprese en función de las varianzas y covarianzas muestrales. Si llamamos “p” al número de variables observadas o indicadores, cuando $\frac{1}{2}[p*(p+1)]$ es mayor o igual al número de parámetros a estimar, no hay seguridad sobre la identificación del modelo. Existen una serie de reglas generales aplicables para identificar un modelo, una de ellas es la regla de los grados de libertad, obtenidos como la diferencia entre el número de varianzas y covarianzas (ecuaciones) y el número de parámetros a estimar. Es una condición necesaria pero no suficiente. Cuando $g < 0$, serán *modelos infraidentificados*, cuando $g = 0$, los modelos son *posiblemente identificados*, y cuando $g > 0$ el modelo está *sobreidentificado*. Otra regla que es condición suficiente pero no necesaria, es que si el modelo es recursivo está identificado, siendo este tipo de modelos aquellos que no contienen efectos circulares o recíprocos entre sus variables. (Casas 2014:5)

- c. Estimación de los parámetros. En esta fase se determinan los valores de los parámetros desconocidos, así como su respectivo error de medición para lo que se utilizan diversos programas computacionales.

Partiendo de que el modelo está identificado, cada uno de los parámetros tendrá un valor único. Si conociésemos los valores de los parámetros del modelo correcto y las varianzas y covarianzas poblacionales, entonces cada elemento de esta matriz sería idéntico al reproducido en la matriz del modelo, pero como la poblacional no es conocida la aproximamos por la matriz de varianzas y covarianzas muestral. El proceso de estimación consiste en la obtención de aquellos valores p de los

parámetros π que ajusten lo mejor posible a la matriz observada, por la que aquellos reproducen. La estimación de coeficientes se realiza mediante procedimientos iterativos de *minimización de desviaciones*, bajo la hipótesis de que nuestro modelo es correcto. Tras la fase de estimación, los tests de bondad del ajuste nos permitirán decidir si la falta de identidad entre la matriz de varianzas y covarianzas muestral y la generada por el modelo, se debe al azar o a la inadecuación del modelo (Casas 2014:5).

Se pueden emplear diferentes funciones de ajuste entre las matrices implicada y observada, aunque todas siguen una estructura similar, la expresión genérica que se minimiza es del tipo. (Batista, J. y Coenders, G. 2000:73) $F = (S - \Sigma(p))' W (S - \Sigma(p))$, en donde S es la matriz observada, $\Sigma(p)$ la matriz implicada, $(S - \Sigma(p))$ son los vectores de residuos y W es la matriz de ponderación.

Según el método de estimación mínimos cuadrados no ponderados, la matriz de ponderación es la matriz identidad, $W = I$. Si empleamos el procedimiento de mínimos cuadrados ponderados bajo normalidad la matriz de ponderación es la inversa de la matriz observada $W = (S \otimes S)^{-1}$. Otro criterio seguido es el de máxima verosimilitud, bajo el supuesto de normalidad multivariante, en donde la matriz de ponderación es la inversa de la matriz implicada, $W = (S(p) \otimes \Sigma(p))^{-1}$. En el caso del método de distribución libre asintótica, la matriz de ponderación es la inversa de una función de los momentos de cuarto orden de las variables observables.

- d. La etapa de diagnóstico de la bondad del ajuste se refiere a la exactitud de los supuestos del modelo especificado para determinar si el modelo es correcto y sirve como aproximación al fenómeno real precisando así su poder de predicción. Si el modelo es correcto y la muestra suficientemente grande, existe una transformación del mínimo de la función de ajuste, llamada *estadístico χ^2 de bondad del ajuste*, que sigue una distribución chi-cuadrado con los mismos grados de libertad g que el modelo. La hipótesis nula a contrastar es que el modelo es bueno, y cuanto mayor sea el valor obtenido del estadístico χ^2 en comparación con los grados de libertad, peor será el ajuste.

Las técnicas de evaluación del modelo pueden ceñirse a una valoración global de la bondad del ajuste o extenderse al análisis detallado de los parámetros y residuos del modelo, con el objetivo de determinar si se han impuesto las restricciones necesarias

al modelo, y si las estimaciones de los parámetros son susceptibles de interpretación plausible y útil para el investigador. A partir del estadístico chi-cuadrado podemos obtener otros indicadores de bondad de ajuste que comparan el valor obtenido de χ^2 para el modelo, con el del modelo base que supone la no-asociación entre las variables del modelo (caso peor). Entre estas medidas encontramos el índice de ajuste normado (NFI), el índice de ajuste no normado (NNFI) y el índice de no centralidad relativo (RNI). Si realizamos un diagnóstico detallado, podemos emplear otros contrastes para ver la significación de parámetros adicionales, como por ejemplo el test de razón de verosimilitud, el test de los multiplicadores de Lagrange, test de Wald, etc. (Casas 2014:6)

- e. La reespecificación del modelo ayuda al investigador a saber si el primer modelo obtenido es el mejor, para lo que es necesario buscar métodos para mejorar el ajuste del mismo añadiendo o eliminando los parámetros estimados del modelo original, con sus justificaciones correspondientes. Para tal caso, el valor del índice de modificación corresponde a la reducción del valor de chi-cuadrado, el cual se sugiere en un mínimo de 3,84 para ser significativa.
- f. La interpretación de los datos ayuda al investigador a establecer el modelo correcto y la aceptación o rechazo de las hipótesis, concluyendo con su investigación.

1.8. Técnicas estadísticas para el análisis de Modelos de Ecuaciones estructurales SEM

Para la estimación de SEM se tienen dos enfoques:

1.8.1. SEM basados en covarianzas

El enfoque basado en covarianzas (CBM) desarrollada alrededor de Karl Jöreskog, es considerado una generalización del Modelo de Senderos, Análisis de Componentes Principales y Análisis de Factores.

1.8.2. SEM basados en componentes.

A este último también se le conoce como basado en varianzas (también conocido como componentes), desarrollado alrededor de Herman Wold bajo el nombre de PLS (Partial Least Squares) puede ser considerado como una generalización del Análisis de Componentes Principales (Tenenhaus 2008: 1).

Ambos enfoques buscan la estimación de los parámetros del modelo, pero los objetivos son diferentes. Mientras que para el enfoque basado en covarianzas CBM, el objetivo es explicativo (busca poner a prueba un modelo), para PLS es predictivo. Ambos objetivos, explicativo y predictivo, son similares a estos objetivos contemplados en la regresión.

1.8.3. Comparativo entre las principales técnicas estadísticas de SEM

Tabla 1 Diferencia entre PLS y CBM

PLS	Métodos CB - SEM
Objetivo: Se adapta mejor para las aplicaciones predictivas y desarrollo de teorías (análisis exploratorio), aunque también puede ser utilizada en la confirmación de teorías (análisis confirmatorio).	Objetivo: Adecuados para confirmación de teorías (análisis confirmatorio).
Basado en la varianza	Basados de la covarianza
Se orienta principalmente para el análisis causal predictivo en situaciones de alta complejidad con un conocimiento teórico escaso.	Más aplicables en situaciones en las que la teoría previa es sólida y tiene como meta un mayor desarrollo y evaluación de la teoría (análisis confirmatorio).
Aplica una secuencia iterativa de OLS (mínimos cuadrados ordinarios) y regresiones lineales múltiples, analizando un constructo cada vez.	Aplican derivadas de segundo orden, como las funciones de máxima verosimilitud para maximizar las estimaciones de los parámetros.
Maximiza la varianza explicada de las variables endógenas latentes. Estima los parámetros del modelo que minimizan la varianza residual de todas las variables dependientes del modelo.	Estima los parámetros del modelo que minimizan la discrepancia entre la matriz de covarianzas estimadas y la muestra.

...continuación

Se adapta particularmente al análisis de muestras de datos pequeñas y datos que no muestra necesariamente la distribución normal multivariante.	Requieren una muestra mínima de 100 o 150, y exigen una distribución normal de los datos.
Análisis de medidas reflectivas y formativas.	Análisis de medidas reflectivas.

Fuente Viera (2010)

II. LOS MODELOS DE ECUACIONES ESTRUCTURALES APLICANDO LA TÉCNICA DE MÍNIMOS CUADRADOS PARCIALES PLS

Los mínimos cuadrados parciales es una técnica de análisis causal diseñada por Wold (1974,1982,1985) para el análisis de datos multidimensionales, y que desde su creación ha sufrido diversas mejoras y modificaciones. PLS fue desarrollado por el profesor sueco Herman Wold (mentor de Karl Jöreskog, fundador de SEM) (Esteves, Pastor y Casanovas, 2002). Originalmente, se denominó NIPALS (Nonlinear Iterative Partial Least Squares) y posteriormente PLS; no obstante, su diseño básico terminó de completarse en 1977 (Cepeda y Roldán, 2004). Esta técnica busca maximizar la varianza explicada de las variables latentes o constructos, al contrario que las técnicas de covarianzas CBM, que busca reproducir la matriz teórica de covarianzas. Aunque es mucho menos popular que las técnicas CBM, PLS se aplica cada vez más, como demuestra el hecho de que en los últimos años se hayan publicado numerosos estudios usando esta técnica en las principales revistas de marketing (Henseler 2009:284). Si se utiliza apropiadamente, PLS tiene ventajas sobre CBM.

2.1. Principales características del PLS

Entre las razones que nos han llevado a aplicar esta técnica para el presente estudio, podemos destacar las siguientes características. (Henseler 2009:283)

- Ofrece valores para las variables latentes, que representan constructos, y que se miden a través de uno o varios indicadores (variables medibles).
- Permite estimar modelos muy complejos con muchas variables latentes y medibles.
- Se basa en supuestos menos estrictos que CBM sobre la distribución de las variables.
- Sirve tanto para modelos reflectivos como formativos de medición.

El objetivo perseguido por la modelización PLS es la predicción de las variables dependientes (tanto latentes como manifiestas). Esta meta se traduce en un intento por maximizar la varianza explicada de las variables independientes, lo que lleva a que las estimaciones de los parámetros estén basadas en la capacidad de minimizar las variables residuales de las variables endógenas Cepeda y Roldán (2004); por su parte, Wold (1979) indica que el objetivo de PLS está diseñado para reflejar las condiciones teóricas y empíricas de las ciencias sociales y del comportamiento, donde son habituales las situaciones con teorías no suficientemente asentadas y escasa información disponible.

2.2. Principales ventajas de PLS

- PLS no precisa que los datos provengan de distribuciones normales o conocidas, no implica ningún modelo estadístico y por tanto, evita la necesidad de realizar suposiciones con respecto a la escala de medida Roldán (2004), en este contexto, las variables pueden estar medidas por diversos niveles o tipos (escalas categóricas, ordinales, de intervalo o ratios).
- Analiza modelos estructurales de ecuación que involucran constructos multiítems con dirección directa e indirecta.
- Puede ser usado para estimar modelos complejos y largos y estimar errores estándar vía procedimiento de remuestreo (Chin, Marcolin y Newsted, 2003:1989-217).
- Barclay, Higgins y Thompson (1995:285-309) concluyen que PLS se recomienda generalmente en modelos de investigación predictivos donde el énfasis se coloca en el desarrollo de una teoría naciente.
- Es un enfoque poderoso para analizar modelos por las demandas mínimas de escalas de medida, tamaño de la muestra y distribuciones residuales (Wold, 1985:581-591).
- Para el tamaño de la muestra, las recomendaciones mínimas están entre 30 y 100 casos, haciéndose más consistente con muestras grandes (Barclay et al. 1995).
- Evitan dos problemas serios, las soluciones inadmisibles y la indeterminación del factor (Fornell y Bookstein, 1982:440-445).

- Es un método general para la estimación de modelos de ruta que involucran constructos latentes medidos indirectamente por los múltiples indicadores (Mathieson y Chin, 2001:89-112).
- Similar a la regresión, pero simultáneamente modela la ruta estructural y el de medición (Chin, Marcolin y Newsted, 2003:189-217).
- Permite el uso de medidas reflectivas y formativas (Mathieson y Chin, 2001:86-112).
- Es aplicable cuando se busca la información y la teoría está menos desarrollada (Fornell y Bookstein, 1982:440-445).
- Trata cada indicador en forma separada permitiendo que cada uno de ellos difiera en el monto de influencia en la estimación del constructo (Chin, Marcolin y Newsted, 2003:189-217), por lo tanto, indicadores con relaciones pobres tienen pesos bajos.
- Se abandona la idea de causalidad y se reemplaza por el concepto de predictibilidad (Cepeda y Roldán, 2004). Mientras que la causalidad garantiza la capacidad de controlar los acontecimientos, la predictividad permite sólo un limitado grado de control.
- (Chin, Marcolin y Newsted 2003:189-217), demostraron que las estimaciones de PLS tienden hacia los auténticos parámetros de la población a medida que se incrementa el número de indicadores y el tamaño de la muestra.

III. METODOLOGÍA

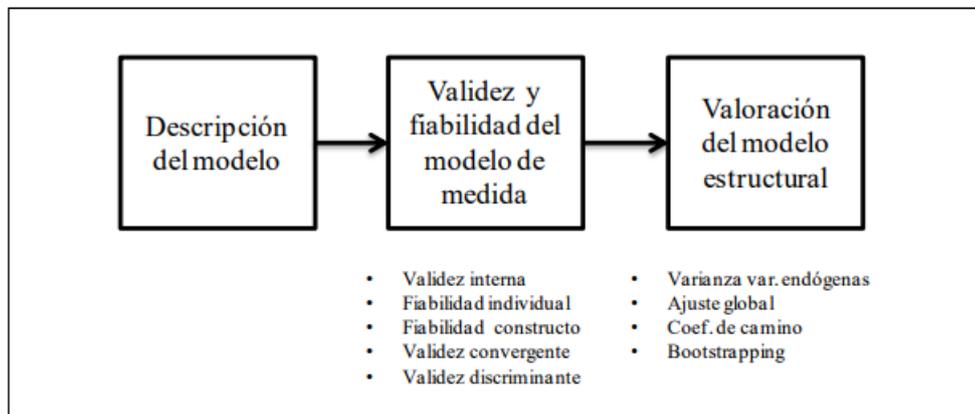
Para utilizar los modelos de ecuaciones estructurales por el método de mínimos cuadrados parciales hay que asegurarse de aplicar el proceso de calidad y refinamiento al grupo de datos, y contar con un modelo conceptual de la investigación.

Tal como indica la Figura 8, la metodología propuesta se basa en tres pasos:

1. Descripción del modelo
2. Validez y fiabilidad del modelo de medida; y
3. Valoración del modelo estructural.

A continuación, se describen cada una de los pasos:

Figura 8: Metodología Propuesta



Fuente: Ramirez *et.al.* (2014)

Los modelos de ecuaciones estructurales se componen de tres pasos:

Paso 1: Descripción del modelo

El modelo causal es descrito gráficamente. El primer paso considera la descripción gráfica del modelo (Barclay et al.1995). Por un lado se debe especificar el modelo estructural, es decir, las relaciones causales entre las variables del modelo, y por otro, las relaciones entre indicadores y constructos, estas relaciones son denominadas modelo de medida.

El modelo de medida es un paso fundamental en este paso, en él se deben identificar las variables latentes (VL) y como éstas se conforman por indicadores correspondientes a variables observables. Existen dos tipos de VL. Una VL se conforma de indicadores reflectivos si estos indicadores son manifestaciones del constructo que representan, luego la VL precede al indicador en un sentido causal, y el indicador está en función de este constructo como indicador reflejo. En cambio, una VL se conforma de indicadores formativos cuando el constructo es expresado como una función de estos indicadores.

Paso 2: Validez y fiabilidad del modelo de medida

En el cual se analizan las cargas factoriales de las variables observables (indicadores) con relación a sus correspondientes variables latentes (constructos). En esta estructura se evalúa la fiabilidad y validez de las medidas del modelo teórico.

Validación del Modelo de Medida

El modelo de medida trata de analizar si los conceptos están medidos correctamente a través de las variables observadas. Este análisis se realiza respecto a los atributos de validez (mide realmente lo que se desea medir) y fiabilidad (lo hace de una forma estable y consistente). Implica el análisis de (Medina y Chaparro, 2008:44-52):

- i. Validación de la Fiabilidad del Ítem. Un ítem para ser considerado en la composición de la variables, debe cargar al menos con 0.5 en el factor, no exceder el 0.3 en otro factor y adquirir un valor agregado para la confiabilidad de la variable (Leidner *et al.*, 1999:633-658); otros investigadores como Hair *et al.* (1999) creen que en la determinación de la importancia y significancia relativa del factor de carga, más del 0.5 es una carga importante pero particularmente por el valor mínimo dicho por Nunnally (1978) que determina que el mínimo es 0.7, y aún más estricto 0.8 para investigación básica, aunque diversos investigadores opinan que esta regla no debería ser tan rígida en las etapas iniciales de desarrollo de escalas

(Chin 998:7). En este sentido, Cepeda y Roldán (2004) quienes consideran que la fiabilidad individual del ítem es valorada examinando las cargas (λ) o correlaciones simples; para aceptar un indicador ha de poseer una carga igual o superior a 0.707 (λ^2 , 50% de la varianza es explicada). Esto implica que la varianza compartida entre el constructo y sus indicadores es mayor que la varianza del error (Sha'ri y Aspinwall, 2000:448-462). Lo anterior es para los indicadores reflectivos; para los formativos, el peso indica la relación al constructo que forman, pero deben de contar con una significancia estadística aceptable de al menos $p < 0.05$ y con un 95% de confianza.

- ii. Consistencia Interna (Validación de la Confiabilidad de Constructos). Establecer niveles aceptables de la validación de los constructos es crítico cuando se usa SEM, lo que dirige a reflexionar si realmente mide lo que se quiere obtener o medir (Straub y Carlson, 1989:147-169). La calidad de la medición del constructo puede ser hecho por medio de su confiabilidad y validez (Im y Grover, 2004:44), y es establecido por el grado en el cual una medida confirma una red de hipótesis relacionadas, generadas en base a la teoría de los conceptos. Se establece durante el análisis estadístico de los datos Zikmund (2003). La confiabilidad de un constructo permite comprobar la consistencia interna de todos los indicadores al medir el concepto, es decir, se evalúa con qué rigurosidad están midiendo las variables manifiestas la misma variable latente Roldán (2004) y dirige a la pregunta de si los constructos verdaderamente son reales (Leidner *et al.*, 1999:633-658). La confiabilidad (estabilidad de la medición) es evaluada por el alfa de Cronbach (0.7), en este caso se usa el estadístico de Fornell y Larcker (1981:39-50) quienes argumentan que su medida es superior a la de Cronbach por los resultados que han obtenido en sus investigaciones (valor mínimo de 0.707).
- iii. Validación Convergente. Determina si los diferentes ítems destinados a medir un concepto o constructo miden realmente lo mismo, entonces el ajuste de dichos ítems será significativo y estarán altamente correlacionados (Barclay, Higgins y Thompson, 1995:285-309). Esta evaluación se lleva a cabo por medio de la Varianza Extraída Media (AVE, por sus siglas en inglés de Average Variance Extracted) que mide el monto de varianza que un constructo captura de sus indicadores relativos a la varianza contenida en el error de medición y debiera ser

más grande que el cuadrado de las correlaciones entre los constructos. Esto es, el monto de varianza compartida entre las variables latentes. Este estadístico puede ser interpretado como una medida de confiabilidad del constructo y como una medición de la evaluación de la validez discriminante (Fornell, 1982:440-445). Los valores de AVE deben de ser mayores de 0.50 o el valor de t student (statistic) significativo con lo que se establece que más del 50% de la varianza del constructo es debido a sus indicadores (Fornell y Larcker, 1981:39). Solo se puede aplicar a bloques dirigidos externamente (Chin, 1998:vii).

- iv. Validación Discriminante. Indica en qué medida un constructo dado es diferente de otros en un modelo de investigación (Barclay, Higgins y Thompson, 1995:285-309). Para que exista la validez discriminante en un constructo han de existir correlaciones débiles entre éste y otras variables latentes que midan fenómenos diferentes. Para esta valoración se usa la raíz cuadrada de AVE Fornell y Larcker (1981); Cepeda y Roldán (2004), la cual debe ser mayor que la varianza compartida entre el constructo y otros constructos en el modelo, la matriz correspondiente proporciona estos valores; es decir, la validez discriminante es evaluada por las correlaciones entre los constructos a un nivel 0.1 (Segars y Grover, 1993:517) y una evaluación alternativa puede ser obtenida comparando AVE con el cuadrado de la correlación entre constructos (Torkzadeh *et al.* 2005)

Paso 3: Valoración del modelo estructural

En el cual se analizan las relaciones de causalidad entre las variables latentes independientes y dependientes.

Validación del Modelo Estructural

El modelo estructural evalúa el peso y la magnitud de las relaciones entre las distintas variables. Para esta valoración se usan dos índices básicos, como lo son los coeficientes de ruta estandarizados (β) y el de la varianza explicada (R^2):

- Representa los coeficientes de ruta estandarizados o pesos de regresión estandarizados, siendo identificado en el nomograma por medio de las flechas que vinculan a los constructos en el modelo interno, este coeficiente se obtiene de la forma tradicional. Chin (1998:viii) propone que para ser considerados significativos,

los coeficientes de ruta estandarizados deberían alcanzar al menos un valor de 0.2 e idealmente situarse por encima de 0.3.

- R^2 indica la varianza explicada por el constructo dentro del modelo. En las variables endógenas debería ser mayor o igual a 0.1, porque valores menores, aún siendo significativos, proporcionan poca información Falk y Miller (1992). Por otra parte, se pueden explorar los cambios en el indicador R^2 para determinar si la influencia de una variable latente particular sobre un constructo dependiente tiene un impacto sustantivo Chin (1988), este último investigador junto con Sellin (1995) indican que para determinar el éxito de la relevancia predictiva de la varianza explicada se utiliza el estadístico de Stone – Geisser test conocido comúnmente como Q^2 , el cual debe ser mayor a 0.0 (cero punto cero) para que sea significativo.

IV. CASO EXPLICATIVO

4.1. Ejemplo

Para la aplicación metodológica se usó la propuesta de Ramirez et al. (2014: 5-12); Venkatesh et al. (2003) formularon el modelo UTAUT (unified theory of acceptance and use of technology - teoría unificada de aceptación y uso de tecnología). Tal como muestra la figura 9, UTAUT propone que la intención de comportamiento (IC) de una tecnología de información es explicada por la expectativa de rendimiento (ER), la expectativa de esfuerzo (EE), y además, el comportamiento de uso (CU) de esta tecnología se explica por la habilidad informacional (HI).

- Se realizó un estudio empírico el segundo semestre de 2013 sobre una muestra correspondiente a 300 alumnos universitarios chilenos con acceso a bases de datos científicas. Para los cálculos de PLS se utilizó el software WarpPLS 4.0.

Para la aplicación se propone el siguiente modelo para analizar:

La figura muestra el modelo de investigación propuesto. A continuación, y basados en la discusión previa y en Venkatesh et al. (2003) se presentan las hipótesis del estudio:

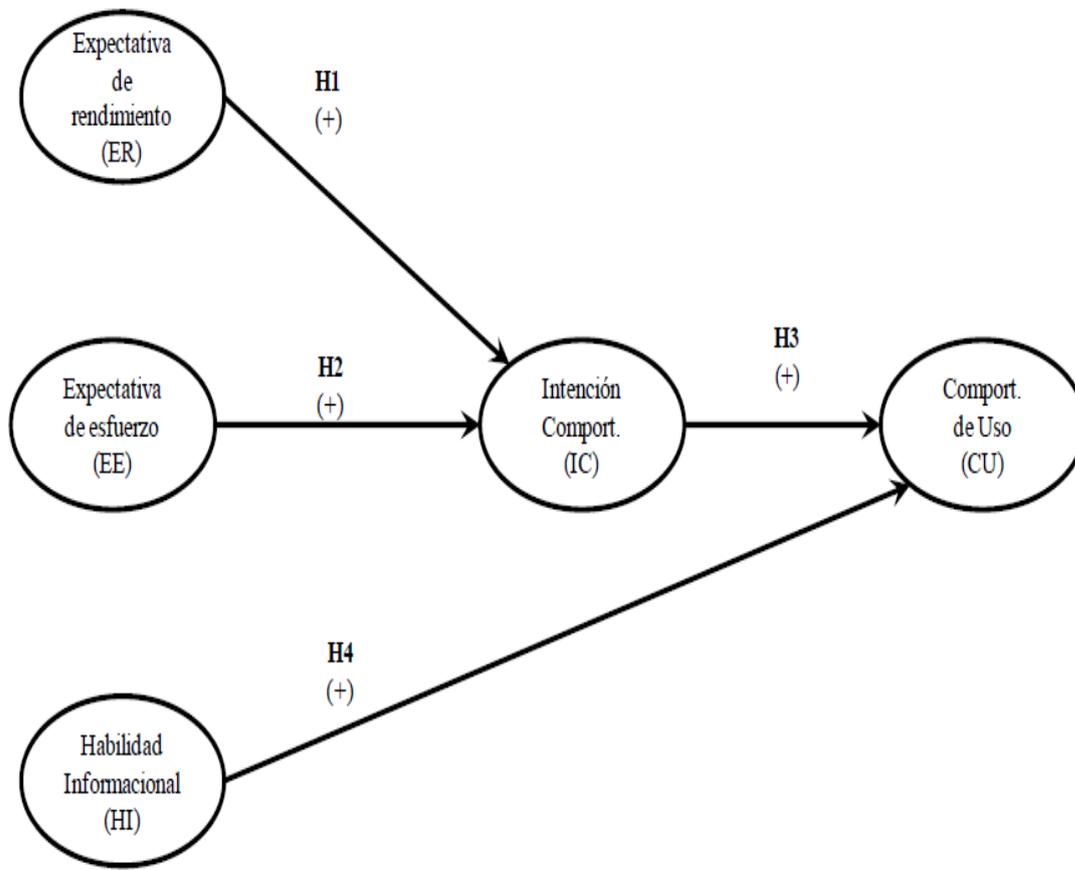
H1: ER se relaciona positivamente con la IC en la adopción de base de datos científicas.

H2: EE se relaciona positivamente con la IC en la adopción de base de datos científicas.

H3: IC se relaciona positivamente con CU en la adopción de base de datos científicas.

H4: HI se relaciona positivamente con CU en la adopción de base de datos científicas.

Figura 9: Modelo propuesto e hipótesis



4.2. Aplicación de la metodología y resultados

Para aplico el programa WarpPLS 4.0 como herramienta para construcción y evaluación del modelo de investigación.

Paso 1: Descripción del modelo. El modelo causal es descrito gráficamente en la Figura 9. En este modelo existen cuatro Variables latentes (VL) conformadas de indicadores reflectivos: ER, EE, IC, y CU. Además, existe una VL conformada de indicadores formativos: HI. Los ítems de cada VL se presentan en el Anexo.

Paso 2: Validez y fiabilidad del modelo de medida. La tabla 2 indica los VIF para cada indicador de HI (todos < 5).

Por otra parte, en la tabla 3 se presenta el VIF de la variable HI (= 1,84), dados estos valores se acepta la validez interna de HI. Además, en la tabla 2 se puede observar que

todas las cargas/pesos de los indicadores de las VL son mayores a 0,7, por lo que se acepta la fiabilidad individual. En la tabla 3 se indican los valores CR, CA, y AVE para la VL conformadas de indicadores reflectivos. Dados estos valores se acepta la fiabilidad de los constructos y su validez convergente. En la tabla 4 se presenta la matriz que justifica la validez discriminante del modelo.

Tabla 2: Cargas y cargas cruzadas de las variables latentes

Indicador	CU	EE	ER	HI	IC	<i>p</i>	VIF
CU1	0,80	-0,23	0,00	-0,08	0,84	***	1,08
CU2	0,80	0,23	0,00	0,08	-0,84	***	1,08
EE1	0,01	0,92	0,03	-0,03	0,06	***	3,11
EE2	0,03	0,93	0,01	0,05	-0,12	***	3,66
EE3	-0,04	0,94	-0,03	-0,02	0,06	***	3,82
ER1	0,04	0,12	0,90	-0,10	-0,10	***	3,20
ER2	0,02	-0,03	0,93	0,02	-0,05	***	3,96
ER3	-0,01	-0,08	0,92	0,02	0,02	***	3,71
ER4	-0,05	-0,01	0,93	0,06	0,12	***	4,20
HI1	0,22	-0,14	0,57	0,72	-0,40	***	2,65
HI2	0,19	-0,02	0,50	0,80	-0,45	***	3,79
HI3	0,12	0,01	0,34	0,79	-0,35	***	3,50
HI4	0,07	-0,16	0,23	0,80	-0,21	***	3,57
HI5	0,12	-0,08	0,16	0,83	-0,24	***	3,91
HI6	0,01	-0,04	0,10	0,84	-0,10	***	3,61
HI7	0,01	-0,05	-0,04	0,89	0,04	***	4,79
HI8	-0,14	-0,01	-0,26	0,84	0,33	***	3,70
HI9	-0,18	0,18	-0,40	0,83	0,29	***	4,08
HI10	-0,13	0,06	-0,37	0,83	0,40	***	4,98
HI11	-0,16	0,09	-0,43	0,79	0,37	***	4,64
HI12	-0,09	0,15	-0,32	0,77	0,27	***	3,37
IC1	0,02	-0,21	0,05	-0,05	0,92	***	2,95
IC2	-0,10	0,00	-0,02	0,02	0,92	***	3,15
IC3	0,07	0,22	-0,03	0,03	0,92	***	2,99

Tabla 3 Coeficientes del modelo de medida

LV	VIF	CR	CA	AVE
Comportamiento de uso (CU)	1,89	0,78	0,43	0,64
Expectativa de rendimiento (ER)	3,52	0,96	0,94	0,85
Expectativa de esfuerzo (EE)	4,46	0,95	0,92	0,87
Habilidad informacional (HI)	1,84	Formativa		
Intención de comportamiento (IC)	3,73	0,94	0,91	0,85

Tabla 4 Análisis de validez discriminante

	CU	EE	ER	HI	IC
CU	0,80				
EE	0,55	0,93			
ER	0,50	0,82	0,92		
HI	0,44	0,59	0,66	0,81	
IC	0,68	0,81	0,70	0,51	0,92

Paso 3: Valoración del modelo estructural. La tabla 5 presenta la varianza de las variables endógenas CU e IC, sus R^2 son 0,47 y 0,66, respectivamente. En la tabla 6 se presentan una serie de índices de ajuste, entre ellos GoF tiene el valor de 0,659, valor que supera el umbral requerido. En la Figura 10 y tabla 5 se presentan los valores de los coeficientes de camino (β) y el valor de significación estadística obtenido a través de Bootstrapping. Como se resume en la tabla 7, estos valores permiten soportar H1 y H3, y soportar solo parcialmente a H2 y H4, dado que sus coeficientes de camino son significativos, pero están entre 0,1 y 0,2.

Figura 10: Resultados del modelo de ecuaciones estructurales

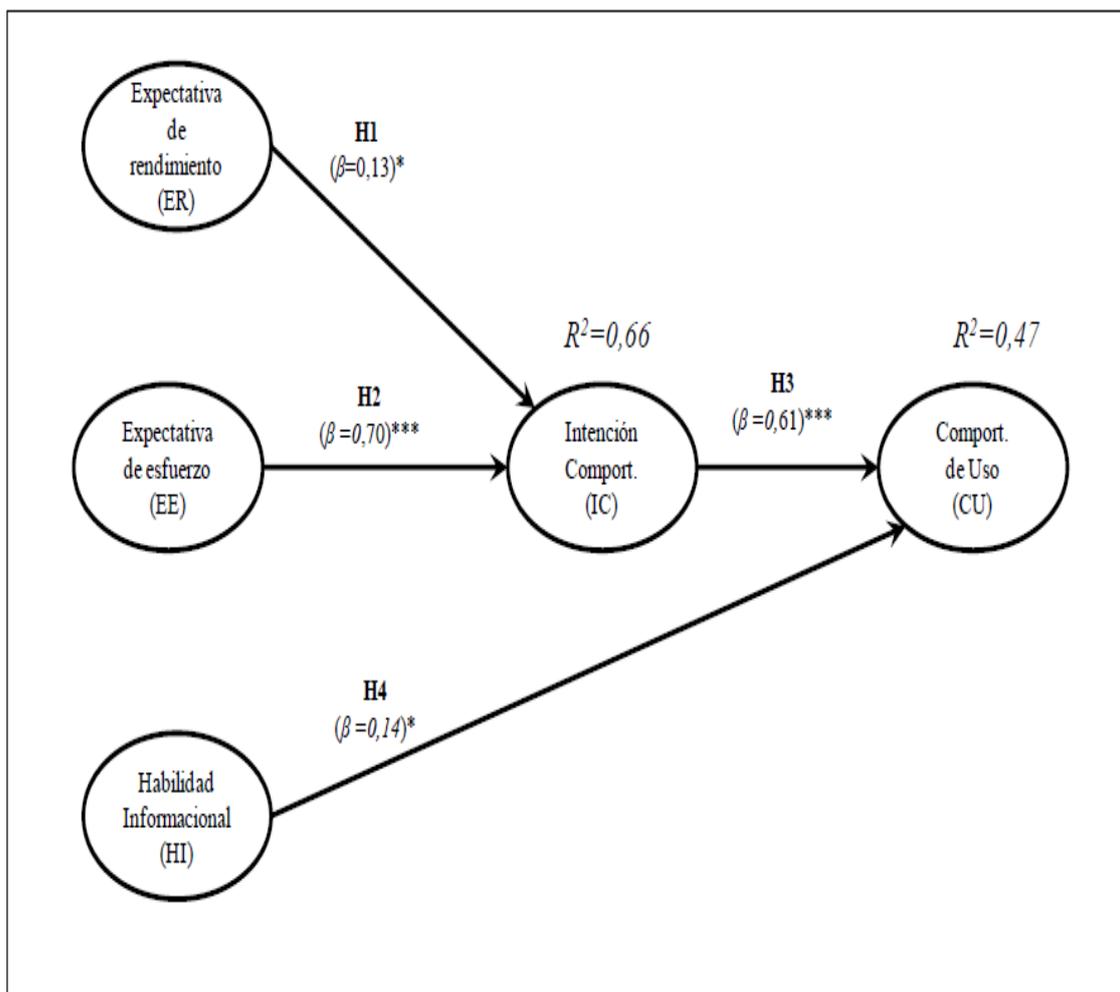


Tabla 5 Resultados del modelo estructural

Variable Dependiente: Comportamiento de uso			ES
R2	0,47		
Q2	0,47		
Habilidad informacional (HI)	0,14	*	0,06
Intención de comportamiento (IC)	0,61	***	0,41
Variable Dependiente: Intención de comportamiento			ES
R2	0,66		
Q2	0,66		
Expectativa de rendimiento (ER)	0,13	*	0,09
Expectativa de esfuerzo (EE)	0,70	***	0,57

Tabla 6 Índices de ajuste del modelo

Index	Value	
Average of path coefficient (APC)	0,393	***
Average of square root (ARS)	0,563	***
Average of variance inflation factor (AVIF)	2,164	
Goodness of fit (GoF)	0,659	

Tabla 7 Resumen de hallazgos

Hipótesis	Variable Independiente	Variable Dependiente	Resultado
H1	Expectativa de rendimiento	Intención de comportamiento	Soportada
H2	Expectativa de esfuerzo	Intención de comportamiento	Soportada parcialmente
H3	Intención de comportamiento	Comportamiento de uso	Soportada
H4	Habilidad informacional	Comportamiento de uso	Soportada parcialmente

V. CONCLUSIONES

Los modelos de ecuaciones estructurales SEM es una extensión de la regresión múltiple, se aplica esta técnica para encontrar relaciones entre variables observables y no observables llamadas (latentes) para pasar posteriormente a estimar los parámetros. Las técnicas de estimación más aplicadas son por el método de la covarianza CBM y por el método de mínimos cuadrados parciales PLS.

La técnica de mínimos cuadrados parciales PLS es la técnica que busca maximizar la varianza explicada de las variables latentes o constructos. PLS tiene como objetivo la predicción, no es preciso que los datos provengan de una distribución normal y puede aplicarse a estudios de muestras pequeñas, permite estimar modelos muy complejos con muchas variables latentes y medibles.

Los resultados del ejemplo en el análisis de PLS del caso indicaron la buena capacidad predictiva del modelo de investigación, y a su vez, la explicación del análisis logró ejemplificar en forma clara la metodología propuesta.

VI. REFERENCIAS BIBLIOGRÁFICAS

- Barclay, D.; C. Higgins; R. Thompson (1995). "The Partial Least Squares (PLS) Approach to Causal Modeling: Personal Computer Adoption and Use as an Illustration". *Technology Studies. Special Issue on Research Methodology*. (2:2), pp. 285-309
- Barclay, D.; C. Higgins; R. Thompson (1995). "The Partial Least Squares (PLS) Approach to Causal Modeling: Personal Computer Adoption and Use as an Illustration". *Technology Studies. Special Issue on Research Methodology*. (2:2), pp. 285-309
- Casas, M. (2002) "Los modelos de ecuaciones estructurales y su aplicación en el índice europeo de satisfacción al cliente". *Revista de la facultad de economía. Universidad San Pablo. España* pp.
- Cepeda, G.; J.L. Roldán 2004. "Aplicando en la Práctica la Técnica PLS en la Administración de Empresas". Congreso de la ACEDE, sep. 19, 20 y 21, Murcia, España
- Chin, W.W. (1998). "Issues and Opinion on Structural Equation Modeling". *MIS Quarterly*. (22:1), pp. vii-xvi
- Chin, W.W. (1998). "Issues and Opinion on Structural Equation Modeling". *MIS Quarterly*. (22:1), pp. vii-xvi
- Chin, W.W.; B.L. Marcolin; P.R. Newsted (2003). "A Partial Least Squares Latent Variable Modeling Approach for Measuring Interaction Effects: Results from a Monte Carlo Simulation Study and an Electronic-Mail Emotion / Adoption Study". *Information Systems Research*. (14:2), pp. 189-217
- Chin, W.W.; P. A. Todd (1995). "On the Use, Usefulness, and Ease of Use of Structural Equation Modeling in MIS Research: A Note of Caution". *MIS Quarterly*. (19:2), pp. 237-246
- Cupani, M. (2012). "Análisis de Ecuaciones Estructurales: conceptos, etapas de desarrollo y un ejemplo de aplicación". *Faculta de Psicología, Universidad Nacional de Córdoba. Colombia* pp. 186-197.
- Esteves, J. J.A. Pastor; J. Casanovas 2002. "Using the Partial Least Squares (PLS) Method to Establish Critical Success Factor Interdependence in ERP Implementation Project". *Universidad Politécnica de Catalunya. Barcelona, Spain*.

- Falk, R.F.; N.B. Miller (1992). "A Primer for Soft Modeling". Akron, Ohio: The University of Akron Press. U.S.A.
- Fernandez V. 2004. Relaciones encontradas entre las dimensiones de las estructuras organizacionales y los componentes del constructo capacidad de absorción: El caso de empresas ubicadas en el territorio español. Tesis doctoral. Barcelona, España. 218 p.
- Fornell, C. 1982. A Second Generation of Multivariate Analysis: An Overview, en C. Fornell [ed.]: A Second Generation of Multivariate Analysis. New York: Praeger Publishers. 1: 1-21
- Fornell, C.; Bookstein, F.L. 1982 A Comparative Analysis of Two Structural Equation Models: Lisrel and PLS Applied to Market Data", en C. Fornell [ed.]: A Second Generation of Multivariate Analysis. New York: Praeger Publishers.1: 289-324
- Fornell, C.; D.F. Larcker (1981). "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error". Journal of Marketing Research. (18:1), pp. 39-50
- Fornell, C.; D.F. Larcker (1981). "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error". Journal of Marketing Research. (18:1), pp. 39-50
- Fornell, C.; F.L. Bookstein (1982). "Two Structural Equation Models: LISREL and PLS Applied to Consumer Exit-Voice Theory". Journal of Marketing Research. (19:4), pp. 440-445
- Fornell, C.; F.L. Bookstein (1982). "Two Structural Equation Models: LISREL and PLS Applied to Consumer Exit-Voice Theory". Journal of Marketing Research. (19:4), pp. 440-445
- García, M. 2011. Análisis causal con ecuaciones Estructurales de la satisfacción Ciudadana con los servicios Municipales Santiago de Compostela, España. 15:16
- Gomez, M. 2011. Estimación de los modelos de ecuaciones Estructurales, del índice mexicano de la satisfacción del usuario de programas sociales mexicanos, con la Metodología de mínimos cuadrados parciales, Mexico D.F. pp.18-20
- Haenlein, M. & Kaplan, A. M. 2004. A beginner's guide to Partial Least Squares Analysis. Understanding Statistics 3(4): 283-297.
- Hair, J.F. Jr.; R.E. Anderson; R.L. Tatham; W.C. Black (1999). "Análisis Multivariante". 5a. Edición, Prentice Hall Iberia. Madrid, España
- Henlein, M. & Kaplan, A. M. 2004. A beginner's guide to partial least squares analysis. Understanding Statistics, 3 (4):283-297.

Henseler, J., Ringle, C y Sinkovics, R 2009 The use of partial least squares de ruta modeling in international marketing, advances in international Marketing (AIM), vol.20 p. 277-320

Im, K.S.; V. Grover (2004). "The Use of Structural Equation Modeling in IS Research: Review and Recommendations". En: M.E. Whitman y A.B. Woszczyński (Eds.). The Handbook of Information Systems Research, pp. 44-63. Idea Group Publishing. Hershey, PA. U.S.A.

Leidner, D.E.; S. Carlsson; J.J. Elam; M. Corrales (1999). "Mexican and Swedish Managers' Perceptions of the Impact of EIS on Organizational Intelligence, Decision Making, and Structure". Decision Science. (30:3), pp. 633-658

Mathieson, K.; W.W. Chin (2001). "Extending the Technology Acceptance Model: The Influence of Perceived User Resources". The DATA BASE for Advances in Information Systems. (32:3), pp. 86-112

Medina, J.M.; J. Chaparro (2008). "The Impact of the Human Element in the Information Systems Quality for Decision Making and User Satisfaction". Journal of Computer Information Systems. (48:2), pp. 44-52

Nunnally, J.C. (1978). "Psychometric Theory". McGraw Hill Editorial. New York, U.S.A.

Ramírez et al. (2014) Propuesta Metodológica para aplicar modelos de ecuaciones estructurales con PLS: El caso del uso de las bases de datos científicas en estudiantes universitarios Universidad Católica del Norte. Chile pp:3-12

Ruiz M, Pardo A y San Martín R. 2010 Modelos de Ecuaciones Estructurales. Universidad Autónoma de Madrid 31(1). España pp. 34-45

Segars, A.H.; V. Grover (1993). "Re-Examining Perceived Ease of Use and Usefulness: A Confirmatory Factor Analysis". MIS Quarterly. (17:4), pp. 517-525

Sellin, N. (1995). "Partial Least Square Modeling in Research on Educational Achievement". En: W. Bos y R.H. Lehmann (Eds.). Reflections on Educational Achievement, pp. 256-267. Waxmann Munster. New York, U.S.A.

Sha'ri, M.Y.; E.M. Aspinwall (2000). "Critical Success Factors in Small and Medium Enterprises: Survey Results". Total Quality Management & Business Excellence. (11,4-6), pp. 448-462

Straub, D.; C. Carlson (1989). "Validating Instrument in MIS Research". MIS Quarterly. (13:2), pp. 147-169

Tomás, J. M., y Fernández, I. 2016. Un primer curso de Modelos de Ecuaciones estructurales con Mplus. Apuntes. 49 p.

- Tomás, J. M., y Fernández, I. 2016. Un primer curso de Modelos de Ecuaciones estructurales con Mplus. Apuntes. Pp. 50-51
- Torkzadeh, G.; X. Koufteros; W.J. Doll (2005). “Confirmatory Factor Analysis and Factorial Invariance of the Impact of Information Technology Instrument”. Omega. (33:2), pp. 107-118
- Vieira, V. 2010 El marketing de relaciones de las universidades privadas de Brasil: efectos de la orientación al aprendizaje y al mercado, Tesis doctoral, universidad de Alcalá, Alcalá de Henares.
- Wold, H. 1985. “Partial Least Squares”. En: Encyclopedia of Statistical Sciences. S. Kotz y N.L. Johnson (Ed.). Vol. 6. Wiley Editorial. U.S.A. pp. 581-591
- Zikmund, W.G. (2003). “Business Research Methods”. Thomson South-Western Editorial, 7th Edition. Ohio, U.S.A.

ANEXO

Ítems usados para estimar el modelo

Comportamiento de uso (CU)

CU1 Frecuencia de veces que utiliza una base de datos científica

CU2 Tiempo que utilizó una base de dato científica en cada acceso

Expectativa de rendimiento (ER)

ER1 Utilizar alguna base de datos científica me ha permitido encontrar la información/documentación que necesitaba para realizar mis investigaciones o trabajos similares.

ER2 Con una base de datos científica he mejorado la calidad de los contenidos de mis investigaciones o trabajos similares.

ER3 Con una base de datos científica se reduce mi tiempo de búsqueda de información/documentación al realizar una investigación o trabajos similares.

ER4 Considero que una base de datos científica es útil.

Expectativa de esfuerzo (EE)

EE1 Considero que es fácil usar una base de datos científica.

EE2 Considero que sería fácil ser experto en el uso de una base de datos científica.

EE3 Necesito poco tiempo para aprender a usar una base de datos científica.

Habilidad informacional (HI)

HI1 Sé definir y estructurar la información que necesito buscar (ej: temas y subtemas).

HI2 Sé identificar distintos formatos de fuentes potenciales de información (libros, artículos, documentos electrónicos, páginas web, diarios, etc.).

HI3 Sé lo beneficioso que es adquirir información de calidad para la realización de una tarea.

HI4 En el momento de realizar una búsqueda de información/documentación, sé identificar cuando es apropiado utilizar google y cuando utilizar una base de datos científica.

HI5 Sé utilizar operadores booleanos o alguna otra estrategia de búsqueda de información/documentación.

HI6	Sé aplicar criterios iniciales para evaluar y descartar la información y sus fuentes.
HI7	Sé determinar si la formulación inicial de la pregunta de investigación debe ser revisada.
HI8	Sé diferenciar una cita de una referencia bibliográfica.
HI9	Conozco al menos una norma para la realización de citas y referencias bibliográficas (APA, Vancouver, AMA).
HI10	Tengo experiencia en la utilización de normas para la realización de citas y referencias bibliográficas.
HI11	Al realizar una cita o referencia bibliográfica normalizada puedo distinguir con claridad el formato de información que estoy utilizando (libros, artículos, documentos electrónicos, páginas web, diarios, etc.).
HI12	Corto y pego información para la realización de investigaciones o trabajos similares.
<i>Intención de comportamiento (IC)</i>	
IC1	Buscaría información/documentación en alguna base de datos científica para realizar investigaciones o trabajos similares.
IC2	Recomendaría alguna base de datos científica para la realización de investigaciones o trabajos similares.
IC3	Volvería a buscar información/documentación en alguna base de datos científica para la realización de nuevas investigaciones o trabajos similares.

Todos los ítems se midieron en una escala Likert de 5 puntos. La escala oscilaba entre "1-nunca" a "5-varias veces a la semana" para CU1 y entre "1-menos de una hora" a "5-más de 6 horas" para CU2. La escala oscilaba entre "1- estar fuertemente en desacuerdo " a "5-estar fuertemente de acuerdo" para el resto de los ítems.