

# **TESIS DOCTORAL**

**Comparación de la eficiencia en la estimación de parámetros  
entre análisis de estructura de covarianzas, modelo de rutas  
mediante mínimos cuadrados parciales y análisis generalizado  
de componentes estructurales: Un estudio de simulación.**

Edgardo Daniel Miranda Zapata

Director

Miguel Ángel Ruiz Díaz

Co-Director

Sergio Escorial Martín

Departamento de Psicología Social y Metodología

Facultad de Psicología

Universidad Autónoma de Madrid

Madrid, 2012

*“Si buscas resultados distintos,  
no hagas siempre lo mismo”*

Albert Einstein

*“...Un camino de mil pasos  
comienza en un solo paso”*

Lao-Tsé

## AGRADECIMIENTOS

En las siguientes líneas quisiera expresar mi agradecimiento a todos quienes colaboraron de alguna forma, directa o indirectamente, en la realización de la presente tesis doctoral. Entre quienes quisiera destacar a mi tutor Miguel Ángel Ruiz, quien en momentos complejos tuvo una disposición abierta y favorable para guiarme en el desarrollo del presente trabajo, brindándome un lugar y los implementos necesarios para desarrollarlo, apoyándome y tolerando todas las inquietudes que me surgieron en el proceso. A mi co-tutor Sergio Escorial, quien me ayudó a dar estructura al presente trabajo, me apoyó y ayudó a ver aspectos positivos cuando no los encontraba por mí mismo. Además quisiera agradecer a todos quienes me ayudaron a comprender un poco más algunos de los temas implicados en este trabajo, ya sea aportando documentos, libros, programas informáticos, sugerencias, respuestas o formulando preguntas necesarias de ser respondidas para comprender distintos temas involucrados. Gracias a: Fabiola González-Betanzos, Francisco Abad, Gastón Sánchez, Giorgio Russolillo, Gregoria Mateos-Aparicio, Heungsun Hwang, Ignacio Montero, Jörg Henseler, Luis Garrido, Michel Tenenhaus, Rafael San Martín y Vincenzo Esposito-Vinci. Particularmente quisiera agradecer a Ludgerio Espinoza por su tolerancia, ya que cada vez que llegó al despacho, me encontraba utilizando el ordenador y su escritorio, y muy amablemente nunca manifestó problema alguno con ello.

Quisiera dar un especial agradecimiento a mi padre Horacio Miranda, quien me instó a cursar el programa de doctorado en metodología de las ciencias del comportamiento y de la salud de la UAM, simplemente sin sus sugerencias y apoyo no estaría escribiendo estas líneas.

Así también, agradezco a la Comisión Nacional de Investigación Científica y Tecnológica del Gobierno de Chile, por otorgarme la Beca Presidente de la República por el periodo 2008-2012 para la realización de estudios doctorales en el extranjero.

Por último quisiera dar las gracias a todos quienes me apoyaron afectivamente, dándome ánimo y acompañándome durante el periodo de formación doctoral y durante el desarrollo de esta tesis doctoral.

A mi familia, particularmente a mi abuela María “Mariji” Contreras, mi madre Magaly Zapata, mis tíos Patricia, Pedro y Cesar Zapata, y mis primas Pamela y Bárbara Zapata, quienes en cada contacto me dieron su apoyo y ánimo para finalizar este trabajo.

A mis amigos de Chile, Fernando Suárez, Leonardo Serendero y Armando Millar, con quienes a pesar de la distancia siempre tuve la sensación de haber estado compartiendo con ellos el día anterior. A las amistades creadas en Madrid, Wolfrang Cárdenas, Andrea Villagrán, Enrique Riquelme y Octavio Londoño, y a mi amigo Claudio Gaete, quienes me acompañaron en los momentos de distensión y me escucharon hablar sobre esta tesis en más de alguna ocasión.

A mis compañeros de entrenamiento de atletismo, en particular a mis entrenadores Manuel Requena y Ricardo Castaño, quienes me brindaron su orientación, apoyo y compañía, lo que me ayudó a finalizar y mejorar en otro tipo de carreras de fondo.

Y finalmente quisiera dar un especial agradecimiento a mi novia Karina Cárdenas, por su cariño, compañía y sobre todo por su tolerancia, ya que al estar más cerca, me ha tenido que soportar en los momentos de máximo estrés e incertidumbre de todo este proceso. A todos ellos y a aquellos que formaron parte, pero que por alguna razón sus nombres no vienen a mi mente en estos momentos, muchas gracias.

---

## ÍNDICE GENERAL

PARTE TEÓRICA.....	1
INTRODUCCIÓN.....	3
Objetivos.....	7
Hipótesis.....	8
Capítulo 1. Análisis de Rutas (PA).....	11
Capítulo 2. Análisis de Estructuras de Covarianzas (ACOV) .....	19
Capítulo 3. Modelo de Rutas mediante Mínimos Cuadrados Parciales (PLS-PM) ....	35
Capítulo 4. Análisis Generalizado de Componentes Estructurales (GSCA) .....	49
Capítulo 5. Comparación de ACOV, PLS-PM y GSCA .....	55
PARTE EMPÍRICA .....	65
Estudio 1: Eficiencia y sesgo en la recuperación de parámetros, con datos ordinales simulados en un modelo con variables manifiestas reflectivas, del Análisis de Estructura de Covarianza y el Modelo de Rutas mediante Mínimos Cuadrados Parciales.....	67
Introducción.....	67
Método .....	70
Simulación de los datos.....	71
Análisis.....	77
Resultados .....	79
Discusión y Conclusiones del Estudio 1 .....	100
Estudio 2: Eficiencia y sesgo en la recuperación de parámetros, con datos ordinales simulados en un modelo con variables manifiestas formativas y reflectivas, del Análisis de Estructura de Covarianza, el Modelo de Rutas mediante Mínimos Cuadrados Parciales y el Análisis Generalizado de Componentes Estructurales.....	105
Introducción.....	105
Método .....	107
Simulación de los datos.....	109
Análisis.....	115
Resultados .....	118
Discusión y Conclusiones del Estudio 2 .....	145
DISCUSIÓN GENERAL.....	153
CONCLUSIONES .....	159
REFERENCIAS .....	163

## ÍNDICE DE TABLAS

Tabla 1. Notación para variables y parámetros de las tres técnicas de MEE .....	55
Tabla 2. Comparación de las tres técnicas de MEE. ....	62
Tabla 3. Umbrales utilizados para obtener variables categóricas de 5 puntos, con tres niveles de asimetría. ....	73
Tabla 4. Tabla de frecuencias de réplicas que no convergen tras 100 iteraciones*, con ACOV y PLS-PM. ....	80
Tabla 5. Pruebas de efectos intra-sujetos e inter-grupo para la varianza promedio de las estimaciones estandarizadas de los parámetros del Modelo Estructural (A) y del Modelo de Medida (B).....	81
Tabla 6. Pruebas de efectos intra-sujetos e inter-sujetos para el sesgo promedio de las estimaciones estandarizadas de los parámetros del Modelo Estructural (A) y del Modelo de Medida (B). ....	84
Tabla 7. Pruebas de efectos intra-sujetos e inter-grupo para el ECM promedio de las estimaciones estandarizadas de los parámetros del Modelo Estructural (A) y del Modelo de Medida (B). ....	89
Tabla 8. Pruebas de efectos intra-sujetos e inter-grupo para el MAD de las estimaciones estandarizadas de los parámetros del Modelo Estructural (A) y del Modelo de Medida (B). ....	94
Tabla 9. Umbrales utilizados para obtener variables categóricas de 5 puntos, con dos niveles de asimetría. ....	111
Tabla 10. Tabla de frecuencias de réplicas que no convergen tras 100 iteraciones*, con ACOV, PLS-PM y GSCA. ....	118
Tabla 11. Pruebas de efectos intra-sujetos e inter-grupo para la varianza promedio de las estimaciones estandarizadas de los parámetros del Modelo Estructural (A) y del Modelo de Medida Reflectivo (b.1) y el Modelo de Medida Formativo (b.2).....	119
Tabla 12. Pruebas de efectos intra-sujetos e inter-grupo para el sesgo promedio de las estimaciones estandarizadas de los parámetros del modelo estructural (A) y del modelo de medida reflectivo (b.1) y formativo (b.2). ....	124
Tabla 13. Pruebas de efectos intra-sujetos e inter-grupo para el ECM promedio de las estimaciones estandarizadas de los parámetros del Modelo Estructural (A) y del Modelo de Medida Reflectivo (b.1) y el Modelo de Medida Formativo (b.2). ....	132
Tabla 14. Pruebas de efectos intra-sujetos e inter-grupo para el promedio del MAD de las estimaciones estandarizadas de los parámetros del Modelo Estructural (A), del Modelo de Medida Reflectivo (b.1) y del Modelo de Medida Formativo (b.2). ....	138

---

## ÍNDICE DE FIGURAS

Figura 1. Ejemplo de diagrama ilustrando dos efectos (XY) determinados en parte por las mismas causas correlacionadas (BC) (Wright, 1920).....	12
Figura 2. Relaciones entre variables representadas en un diagrama de rutas. ....	14
Figura 3. Relaciones entre variables incorporadas por el ACOV.....	20
Figura 4. Ejemplo de Diagrama de rutas de un Modelo de Rutas mediante Mínimos Cuadrados Parciales. ....	37
Figura 5. Ejemplo de Diagrama de rutas de un Análisis Generalizado de Componentes Estructurales.....	49
Figura 6. Diagrama de rutas del modelo poblacional, con variables manifiestas reflectivas. ....	71
Figura 7. Medias de la varianza promedio de las estimaciones del Modelo Estructural con ACOV y PLS-PM, de acuerdo al nivel de asimetría.....	82
Figura 8. Medias de la varianza promedio de las estimaciones del Modelo Estructural con ACOV y PLS-PM, de acuerdo a la especificación del modelo.....	83
Figura 9. Medias del sesgo promedio de las estimaciones del Modelo Estructural con ACOV y PLS-PM, de acuerdo a la especificación del modelo. ....	85
Figura 10. Medias del sesgo promedio de las estimaciones del Modelo Estructural con ACOV y PLS-PM, de acuerdo al nivel de asimetría.....	86
Figura 11. Medias del sesgo promedio de las estimaciones del Modelo de Medida con ACOV y PLS-PM, de acuerdo al nivel de asimetría.....	87
Figura 12. Medias del sesgo promedio de las estimaciones del Modelo de Medida con ACOV y PLS-PM, de acuerdo al tamaño de la muestra. ....	88
Figura 13. Medias del EMC de las estimaciones del Modelo Estructural con ACOV y PLS-PM, de acuerdo al nivel de asimetría. ....	90
Figura 14. Medias del EMC de las estimaciones del Modelo Estructural con ACOV y PLS-PM, de acuerdo con la especificación del modelo.....	91
Figura 15. Medias del EMC de las estimaciones del Modelo Estructural con ACOV y PLS-PM, de acuerdo el tamaño muestral. ....	92
Figura 16. Promedios del MAD de las estimaciones del Modelo Estructural con ACOV y PLS-PM, de acuerdo a la especificación del modelo, con asimetría nula.....	95

Figura 17. Promedios del MAD de las estimaciones del Modelo Estructural con ACOV y PLS-PM, de acuerdo a la especificación del modelo, con asimetría media. ....	95
Figura 18. Promedios del MAD de las estimaciones del Modelo Estructural con ACOV y PLS-PM, de acuerdo a la especificación del modelo, con asimetría alta. ....	96
Figura 19. Promedios del MAD de las estimaciones del Modelo Estructural con ACOV y PLS-PM, de acuerdo al tamaño de la muestra. ....	97
Figura 20. Promedios para el MAD de las estimaciones del Modelo de Medida con ACOV y PLS-PM, de acuerdo al tamaño de la muestra con asimetría nula. ....	98
Figura 21. Promedio del MAD de las estimaciones del Modelo de Medida con ACOV y PLS-PM, de acuerdo al tamaño de la muestra con asimetría media. ....	98
Figura 22. Promedio del MAD de las estimaciones del Modelo de Medida con ACOV y PLS-PM, de acuerdo al tamaño de la muestra con asimetría alta. ....	99
Figura 23. Diagrama de rutas del modelo poblacional, con variables manifiestas reflectivas y formativas. ....	109
Figura 24. Medias de la varianza promedio de las estimaciones del Modelo Estructural con ACOV, PLS-PM y GSCA, según el tamaño muestral. ....	120
Figura 25. Medias de la varianza promedio de las estimaciones de parámetros del Modelo de Medida Reflectivo con ACOV, PLS-PM y GSCA, en nivel de asimetría nulo, según el tamaño muestral. ....	121
Figura 26. Medias de la varianza promedio de las estimaciones de parámetros del Modelo de Medida Reflectivo con ACOV, PLS-PM y GSCA, en nivel de asimetría alto, según el tamaño muestral. ....	122
Figura 27. Medias de la varianza promedio de las estimaciones de parámetros del Modelo de Medida Formativo con ACOV, PLS-PM y GSCA, en nivel de asimetría nulo, según el tamaño muestral. ....	123
Figura 28. Medias de la varianza promedio de las estimaciones de parámetros del Modelo de Medida Formativo con ACOV, PLS-PM y GSCA, en nivel de asimetría alto, según el tamaño muestral. ....	124
Figura 29. Medias del sesgo promedio de las estimaciones de parámetros del Modelo de Estructural con ACOV, PLS-PM y GSCA, según la especificación del modelo. ....	125
Figura 30. Medias del sesgo promedio de las estimaciones de parámetros del Modelo Estructural con ACOV, PLS-PM y GSCA, según el tamaño de la muestra. ....	127
Figura 31. Medias del sesgo promedio de las estimaciones de parámetros del Modelo de Medida Reflectivo con ACOV, PLS-PM y GSCA, en nivel de asimetría nulo, según el tamaño muestral. ....	128



---

Figura 32. Medias del sesgo promedio de las estimaciones de parámetros del Modelo de Medida Reflectivo con ACOV, PLS-PM y GSCA, en nivel de asimetría alto, según el tamaño muestral. ....	128
Figura 33. Medias del sesgo promedio de las estimaciones de parámetros Modelo de Medida Formativo con ACOV, PLS-PM y GSCA, en nivel de asimetría nulo, según el tamaño muestral. ....	130
Figura 34. Medias del sesgo promedio de las estimaciones de parámetros del Modelo de Medida Formativo con ACOV, PLS-PM y GSCA, en nivel de asimetría alto, según el tamaño muestral. ....	130
Figura 35. Medias del ECM de las estimaciones de parámetros del Modelo Estructural con ACOV, PLS-PM y GSCA, en nivel de asimetría nulo, según el tamaño muestral.....	133
Figura 36. Medias del ECM de las estimaciones de parámetros del Modelo Estructural con ACOV, PLS-PM y GSCA, en nivel de asimetría alto, según el tamaño muestral.....	133
Figura 37. Medias del ECM de las estimaciones de parámetros del Modelo de Medida Reflectivo con ACOV, PLS-PM y GSCA, en nivel de asimetría nulo, según el tamaño muestral.....	135
Figura 38. Medias del ECM de las estimaciones de parámetros del Modelo de Medida Reflectivo con ACOV, PLS-PM y GSCA, en nivel de asimetría alto, según el tamaño muestral.....	135
Figura 39. Medias del ECM de las estimaciones de parámetros del Modelo de Medida Formativo con ACOV, PLS-PM y GSCA, en nivel de asimetría nulo, según el tamaño muestral.....	136
Figura 40. Medias del ECM de las estimaciones de parámetros del Modelo de Medida Formativo con ACOV, PLS-PM y GSCA, en nivel de asimetría alto, según el tamaño muestral.....	137
Figura 41. Promedios del MAD de las estimaciones del Modelo Estructural con ACOV, PLS-PM y GSCA, de acuerdo al tamaño de la muestra, en asimetría nula.....	139
Figura 42. Promedios del MAD de las estimaciones del Modelo Estructural con ACOV, PLS-PM y GSCA, de acuerdo al tamaño de la muestra, en asimetría alta.....	139
Figura 43. Promedios del MAD de las estimaciones del Modelo Estructural con ACOV, PLS-PM y GSCA, de acuerdo a la especificación del modelo.....	140
Figura 44. Promedios del MAD las estimaciones del Modelo de Medida Reflectivo con ACOV, PLS-PM y GSCA, de acuerdo al tamaño de la muestra, en asimetría nula.....	141

Figura 45. Promedios del MAD de las estimaciones del Modelo de Medida Reflectivo con ACOV, PLS-PM y GSCA, de acuerdo al tamaño de la muestra, en asimetría alta. .... 142

Figura 46. Promedios del MAD de las estimaciones del Modelo de Medida Formativo con ACOV, PLS-PM y GSCA, de acuerdo al tamaño de la muestra, en asimetría nula. .... 143

Figura 47. Promedios del MAD de las estimaciones del Modelo de Medida Formativo con ACOV, PLS-PM y GSCA, de acuerdo al tamaño de la muestra, en asimetría alta. .... 143

---

## ANEXOS

Anexo A. Descripción de la sintaxis para la simulación de datos con escala ordinal con Matlab, a partir de la matriz de varianzas y covarianzas del modelo con variables manifiestas reflectivas, modelo Más especificado. ....	179
Anexo B. Matriz de varianzas y covarianzas poblacional, $\Sigma(\theta)$ , del modelo con variables manifiestas reflectivas, Más especificado. ....	183
Anexo C. Descripción de la sintaxis PRELIS para el cálculo de la matriz de correlaciones policóricas. ....	184
Anexo D. Descripción de la sintaxis LISREL para el análisis ACOV del modelo con variables manifiestas reflectivas, Más especificado. ....	185
Anexo E. Descripción de la sintaxis R para el análisis PLS-PM del modelo con variables manifiestas reflectivas, Más especificado. ....	187
Anexo F. Descripción de la sintaxis Visual FoxPro, para la ejecución automatizada del análisis ACOV con LISREL, del modelo con variables manifiestas reflectivas, Más especificado. ....	188
Anexo G. Descripción de la sintaxis Visual FoxPro, para la ejecución automatizada del análisis PLS-PM con R, del modelo con variables manifiestas reflectivas, Más especificado. ....	191
Anexo H. Descripción de la sintaxis Visual FoxPro, para la extracción automatizada de información desde archivos de resultados de LISREL, al analizar el modelo con variables manifiestas reflectivas. ....	195
Anexo I. Descripción de la sintaxis Visual FoxPro, para la extracción automatizada de información desde archivos de resultados de R, al analizar el modelo con variables manifiestas reflectivas. ....	201
Anexo J. Descripción de la sintaxis para la simulación de datos con Matlab, a partir de la generación de variables aleatorias del modelo con variables manifiestas formativas y reflectivas, modelo Más especificado. ....	205
Anexo K. Descripción de la sintaxis LISREL para el análisis ACOV del modelo con variables manifiestas formativas y reflectivas, Más especificado. ....	210
Anexo L. Descripción de la sintaxis R para el análisis PLS-PM del modelo con variables manifiestas formativas y reflectivas, Más especificado. ....	212
Anexo M. Descripción de la sintaxis Visual FoxPro, para la ejecución automatizada de análisis ACOV con LISREL, modelo con variables manifiestas formativas y reflectivas, Más especificado. ....	213
Anexo N. Descripción de la sintaxis Visual FoxPro, para la ejecución automatizada de análisis PLS-PM con R, modelo con variables manifiestas formativas y reflectivas, Más especificado. ....	216

Anexo O. Descripción de la sintaxis Visual FoxPro, para la extracción automatizada de información desde archivos de resultados de LISREL, al analizar el modelo con variables manifiestas formativas y reflectivas. .... 220

Anexo P. Descripción de la sintaxis Visual FoxPro, para la extracción automatizada de información desde archivos de resultados de R, al analizar el modelo con variables manifiestas formativas y reflectivas. .... 225

Anexo Q. Descripción de la sintaxis Visual FoxPro, para la extracción automatizada de información desde archivos de resultados de GeSCA, al analizar el modelo con variables manifiestas formativas y reflectivas. .... 229

## GLOSARIO DE ABREVIATURAS

- ACOV** Análisis de Estructura de Covarianzas. Técnica de análisis de modelos de ecuaciones estructurales, basada en covarianzas. Desarrollada por Karl Jöreskog en 1970, que busca confirmar relaciones entre variables latentes teóricamente asumidas. Se caracteriza por plantear los indicadores definidos como una combinación de variables latentes del modelo y contar con índices globales de bondad de ajuste del modelo a los datos.
- ACOVs** Análisis de Estructura de Covarianzas. Nombre del primer programa informático para la aplicación del Análisis de Estructura de Covarianzas, desarrollado por Karl Jöreskog, Gunnar Gruvaeus y Marielle van Thillo en 1970, posteriormente sustituido por el programa LISREL.
- GeSCA** Análisis Generalizado de Componentes Estructurales. Nombre del programa informático para la aplicación del Análisis Generalizado de Componentes Estructurales, desarrollado por Heungsun Hwang en 2009, disponible para libre acceso en: [www.sem-gesca.org](http://www.sem-gesca.org).
- GSCA** Análisis Generalizado de Componentes Estructurales. Técnica de análisis de modelos de ecuaciones estructurales, basada en componentes. Desarrollada por Heungsun Hwang y Yoshio Takane en 2004. Busca maximizar la varianza explicada de las variables dependientes de un modelo. Se caracteriza por plantear las variables latentes como una combinación ponderada de las variables latentes del modelo y por contar con índices globales de bondad de ajuste del modelo.
- LISREL** Relaciones Lineales Estructurales. Nombre del programa informático para la aplicación del Análisis de Estructura de Covarianzas, originalmente desarrollado por Karl Jöreskog y Marielle van Thillo en 1972. Versiones posteriores fueron desarrolladas por Karl Jöreskog y Dag Sörbom. El programa sigue siendo utilizado y actualizado de manera periódica.
- MEE** Modelos de Ecuaciones Estructurales. Familia de modelos estadísticos multivariantes que permiten estimar el efecto y las relaciones entre múltiples variables a través de un conjunto de ecuaciones simultáneas que

definen el modelo de medida o modelo externo y el modelo estructural o modelo interno.

- PA Análisis de Rutas. Primera metodología para analizar sistemas de ecuaciones múltiples, desarrollada por el biómetra Sewall Wright en 1918. Se caracteriza por utilizar sólo variables observables, estableciendo relaciones entre ellas.
- PLS-PM Modelo de Rutas mediante Mínimos Cuadrados Parciales. Técnica de análisis de modelos de ecuaciones estructurales, basada en componentes. Desarrollada por Herman Wold en 1982(a), que busca maximizar la varianza explicada de las variables dependientes de un modelo. Se caracteriza por plantear las variables latentes como una combinación ponderada de las variables latentes del modelo. No cuenta con índices de bondad global de ajuste del modelo.
- plsmp Nombre del paquete del Proyecto R, desarrollado por Gastón Sánchez y Laura Trinchera en 2012, para la aplicación del Modelo de Rutas mediante Mínimos Cuadrados Parciales, se encuentra disponible para libre acceso en: <http://cran.r-project.org/web/packages/plsmp/index.html>.
- R Proyecto libre, consistente en un ambiente y lenguaje de programación para el análisis estadístico y gráfico, inicialmente desarrollado por Robert Gentleman y Ross Ihaka en 1993. Actualmente es responsabilidad de R Development Core Team y se encuentra disponible para libre acceso en: <http://www.r-project.org/>.
- VFP Visual FoxPro. Lenguaje de programación que provee de herramientas para crear y administrar aplicaciones y componentes de bases de datos. Lanzado al comercio por Microsoft Corporation en 1995. VFP 9.0 es la última versión, la que fue publicada en 2007. Actualmente Microsoft ha decidido no continuar con Microsoft Visual FoxPro. Sin embargo, existe una comunidad de desarrolladores que sigue trabajando en él, tales como PortalFox y Mundo Visual FoxPro.

# PARTE TEÓRICA





## INTRODUCCIÓN

Los Modelos de Ecuaciones Estructurales (MEE) son cada vez más utilizados en distintos ámbitos de las ciencias del comportamiento y de la salud (Gefen, Rigdon y Straub, 2011; Guo, Perron y Gillespie, 2009; Hooper, Coughlan y Mullen, 2008; Urbach y Ahlemann, 2010). Una definición posible de los MEE es que son "...una familia de modelos estadísticos multivariantes que permiten estimar el efecto y las relaciones entre múltiples variables" (Ruiz, Pardo y San Martín, 2010, p.34), de las cuales "se estima el error de medida y los parámetros teóricos se ajustan en consecuencia" (Guo, Perron y Gillespie, 2009, p.1557), admitiendo múltiples variables manifiestas de constructos latentes y la estimación de fiabilidad y validez (Bollen y Long, 1993). Lo que junto con su propósito fundamental de "confirmar mediante el análisis de la muestra (o muestras) las relaciones propuestas a nivel teórico" (Ruiz, 2000, p.12), han favorecido que se hayan convertido en uno de los métodos más utilizados para la validación de instrumentos y el contraste de relaciones entre constructos (Gefen, Straub y Boudreau, 2000), representando un avance importante en la investigación en el área de las ciencias sociales (Guo, Perron y Gillespie, 2009).

Esta definición construida con el aporte de distintos autores incluye varios elementos de los MEE que son ampliamente aceptados y reconocidos como características propias de los MEE. De hecho, varios autores (p.e. Hancock y Mueller, 2006; Kaplan, 2009; Kline, 2011; Mulaik, 2009; Rigdon, 1998) reservan la denominación "Modelos de Ecuaciones Estructurales" para aquella aproximación que cumple con todos los elementos de la definición planteada en el párrafo anterior: la Aproximación Basada en Covarianzas, cuya aplicación habitual es en el ámbito de la psicología. Otros autores (p.e. Anderson y Gerbing, 1988; Chin, 1998a; Fornell y Bookstein, 1982; Hwang, Malhotra, Kim, Tomiuk y Hong, 2010a; Jöreskog y Wold,

1982; Lévy, González y Muñoz, 2006; Tenenhaus, Esposito Vinzi, Chatelin y Lauro, 2005) incluyen dentro de los MEE la Aproximación Basada en Componentes. La que incluye aplicaciones en áreas como la economía, marketing y química, ampliando las opciones frente a la primera.

Estas dos aproximaciones de MEE presentan distintas posibilidades de aplicación o técnicas de análisis. Entre las técnicas de la aproximación basada en covarianzas se encuentra el Análisis de Estructura de Covarianza (ACOV; Jöreskog, 1970a, 1970b, 1979a y 1979b), el método de Skron dal y Laake (SL; Skron dal y Laake, 2001) y el método de Croon (Croon, 2002). Entre las aplicaciones de la aproximación basada en componentes se encuentra el Modelo de Rutas mediante Mínimos Cuadrados Parciales (PLS-PM; Wold, 1977 y 1980) y el Análisis Generalizado de Componentes Estructurales (GSCA; Hwang y Takane, 2004).

Esta tesis doctoral se centrará en las técnicas de MEE: ACOV, PLS-PM y GSCA. Como se puede apreciar en el párrafo anterior ACOV y PLS-PM son las dos técnicas primeramente desarrolladas de MEE, lo que ha favorecido que sean las más utilizadas, estudiadas y comparadas. De hecho, en 1982, Jöreskog y Wold comparan sus propuestas. Cada una de las cuales representa una de las dos aproximaciones de MEE. Estos autores plantean que con ACOV se obtienen estimaciones más robustas. Esto sumado a que, ya en 1972, Jöreskog y van Thillo desarrollaron el programa informático LISREL (Jöreskog y van Thillo, 1972), para la aplicación de ACOV, ha favorecido su mayor popularidad. En contraposición, PLS-PM sólo ha contado con el programa informático LVPLS, desarrollado por Jan-Bernd Lohmöller, desde 1984. En la actualidad ambos tipos de modelado cuentan con diversos programas específicos de análisis o han sido incluidos como rutinas de paquetes de análisis de datos: AMOS (Arbuckle, 1994), CALIS (SAS Institute Inc., 2008), EQS (Bentler, 1995), Mplus (Muthén y Muthén, 1998-2010) y R v. 2.14.0 (R Development core team, 2011) con los

---

paquetes “sem” v. 3.0-0 (Fox, Nie y Byrne, 2012), “lavaan” (Rosseel, 2012) y “OpenMx” (Boker y cols., 2011), para ACOV; y PLS-Graph (Chin, 2001), PLS-GUI (Li, 2003), PLSPATH (Sellin, 1989), R v. 2.14.0 (R Development core team, 2011) con los paquetes “plsrm” (Sánchez y Trinchera, 2012) y “semPLS” (Monecke y Leisch, 2012), SmartPLS (Ringle, Wende y Will, 2005a), SPAD-PLS (Test and Go, 2006), VisualPLS (Fu, 2006) y XLSTAT (Addinsoft, 2009), para PLS-PM.

Cabe señalar que PLS-PM surge, en primera instancia, como una alternativa frente al incumplimiento de los requisitos de tamaño muestral y de distribución normal de ACOV, cuando se utiliza con el método de estimación de parámetros de Máxima Verosimilitud (Maximum Likelihood, ML). Dado que PLS-PM permite obtener estimaciones de los parámetros sin requisitos para la distribución de los datos y con tamaños muestrales pequeños (Reinartz, Haenlein, y Henseler, 2009; Tenenhaus, 2008).

Ringle, Sarstedt y Straub (2012), realizaron una revisión de 65 estudios publicados en *MIS Quarterly*, desde 1992 a 2011, que utilizan PLS-PM. Los resultados son concordantes con los encontrados por Hair, Sarstedt, Ringle y Mena (2012), quienes revisaron 204 estudios que utilizan PLS-PM, publicados en 30 revistas científicas en el periodo 1981-2010. En ambos estudios la mayoría de los fundamentos para la elección de dicha técnica de MEE se encuentran relacionados con la utilización de un tamaño muestral pequeño, datos no normales y el uso de variables formativas. Resulta llamativo que sólo la mitad de los estudios consideren la concordancia del objetivo del estudio con el objetivo de PLS-PM para fundamentar su elección, tal como propuso Wynne Chin, ya en 1995.

Una de las principales limitaciones de PLS-PM frente a ACOV, es que carece de índices de bondad de ajuste globales, lo que limita su aplicación en el ámbito de la

psicología donde la principal aplicación de los MEE se encuentra relacionada con la comprobación en los datos de los modelos planteados teóricamente. Por lo que cobra relevancia la propuesta del GSCA por parte de Hwang y Takane (2004). El GSCA es una técnica de MEE que mantiene las propiedades de PLS-PM y cuenta con índices de bondad de ajuste globales, por lo que constituye una alternativa más adecuada que PLS-PM, frente a ACOV cuando existe incumplimiento de requisitos de distribución de los datos y cuando se utilizan tamaños muestrales pequeños (Hwang y cols., 2010a).

Hwang y cols. (2010a), estudian la recuperación de parámetros de tres técnicas de MEE: ACOV, PLS-PM y GSCA con datos continuos simulados, en un modelo con variables manifiestas reflectivas que incluye efectos cruzados. Los autores encontraron que la única condición determinante en la recuperación de parámetros fue la especificación del modelo. En concreto, cuando el modelo se encontraba correctamente especificado, ACOV recuperó los parámetros sin sesgo y recuperó mejor los parámetros que PLS-PM y GSCA. Sin embargo, cuando el modelo se encontraba incorrectamente especificado GSCA recuperó mejor los parámetros.

Por su parte, Tenenhaus (2008), compara las tres técnicas de MEE utilizando el modelo del Índice Europeo de Satisfacción del Consumidor (ECSI), con datos referidos a proveedores de telefonía móvil. El modelo incluye siete variables latentes, con escala continua para las variables manifiestas (de 0 a 100). El autor informa que ACOV (utilizando el método de estimación de parámetros ULS), PLS-PM y GSCA entregan resultados muy similares. Sin especificar el tamaño muestral utilizado, ni la incorporación de condiciones experimentales.

Tomando en cuenta, la diferencia entre indicadores efecto (reflectivos) e indicadores causa (formativos), en cuanto a que los primeros representan el reflejo o manifestación de un constructo (Fornell y Bookstein, 1982) y los segundos “son

variables observadas que afectan directamente su variable latente” (Bollen, 2002, p. 616). Se considera necesario precisar que en el presente trabajo se incorporan ambos tipos de indicadores en los modelos teóricos utilizados.

Hasta el momento, dos investigaciones han evaluado la recuperación de parámetros de ACOV, PLS-PM y GSCA. Ninguna ha utilizado variables manifiestas en escala ordinal, ni ha utilizado variables manifiestas formativas.

## **Objetivos**

La presente tesis doctoral busca profundizar en el estudio de la recuperación de parámetros. Partiendo del reconocimiento de distintas técnicas de MEE, se pretende:

- 1.- Evaluar la eficiencia y sesgo en la estimación de parámetros de las técnicas de MEE: ACOV, PLS-PM y GSCA, en condiciones que no se han utilizado conjuntamente en estudios previos, tales como la utilización de datos sintéticos simulados con escala ordinal, en un modelo con variables manifiestas reflectivas y formativas. Además de incorporar condiciones habitualmente utilizadas en el estudio de la recuperación de parámetros tales como distintos tamaños muestrales, niveles de asimetría y especificaciones del modelo; con niveles habituales en psicología y educación (Micceri, 1989); y economía (Hulland, Ryan y Rayner, 2005, citados por Hwang y cols., 2010a) y valores extremos que superan las limitaciones de estudios previos (Hwang y cols., 2010a).
  
- 2.- Comparar la precisión en la recuperación de parámetros de las técnicas de MEE: ACOV, PLS-PM y GSCA en las distintas condiciones experimentales señaladas en el objetivo 1.

## Hipótesis

Se plantean las siguientes hipótesis de investigación:

- (1). La eficiencia en la recuperación de parámetros de las aproximaciones basadas en componentes, será similar y mayor que la eficiencia presentada por ACOV en condiciones de tamaño muestral pequeño y asimetría alta.
  
- (2). La eficiencia en la recuperación de parámetros de ACOV será mayor que la presentada por las aproximaciones basadas en componentes, cuando las condiciones experimentales favorezcan el cumplimiento de supuestos de distribución de los datos (asimetría nula) y requisitos de tamaño muestral (tamaño muestral grande y muy grande).
  
- (3). La precisión en la recuperación de parámetros de las aproximaciones basadas en componentes, será similar y mayor que la presentada por ACOV en condiciones de tamaño muestral pequeño y asimetría alta.
  
- (4). La precisión en la estimaciones de parámetros de ACOV será mayor que la presentada por las aproximaciones basadas en componentes, cuando las condiciones experimentales favorezcan el cumplimiento de supuestos de distribución de los datos (asimetría nula) y requisitos de tamaño muestral (tamaño muestral grande y muy grande).

- (5). La precisión en la recuperación de parámetros de GSCA será mayor que la precisión de ACOV y PLS-PM en condición de modelo menos especificado.

La tesis se estructura en dos secciones. La primera parte es puramente teórica e incluye un capítulo para el Análisis de Rutas (Path Analysis, PA), un capítulo para cada una de las tres técnicas de MEE: ACOV, PLS-PM y GSCA, y un capítulo sobre la comparación de las mismas. La segunda parte es empírica e incluye la realización de dos estudios de simulación que buscan responder a los objetivos de esta tesis doctoral. El primero de ellos busca evaluar la eficiencia y sesgo de ACOV y PLS-PM en la recuperación de parámetros con datos simulados con escala ordinal, en un modelo con variables manifiestas reflectivas, considerando distintas condiciones de: a) asimetría, b) especificación del modelo y c) tamaño muestral. Utilizando el método de estimación de parámetros ULS para ACOV, el que se ha demostrado es el método más eficiente cuando se utilizan variables manifiestas con escala ordinal (Forero, Maydeu-Olivares y Gallardo-Pujol, 2009).

El segundo estudio incorpora la técnica de MEE: GSCA e incluye variables manifiestas formativas en la definición del modelo. Así, este segundo estudio busca comparar ACOV, PLS-PM y GSCA, respecto de su eficiencia y sesgo en la recuperación de parámetros en un modelo que incluye variables manifiestas reflectivas y formativas, con datos simulados con escala ordinal, considerando distintos niveles de asimetría, especificación del modelo y tamaño muestral; superando algunas de las limitaciones presentes en el trabajo de Hwang, y cols. (2010a) y señaladas por los propios autores, en cuanto a los niveles de asimetría y tamaño muestral. Este estudio reduce la cantidad de condiciones y de réplicas por condición (en comparación con el primer estudio de la parte práctica) para posibilitar la incorporación de la técnica GSCA, dada la limitación para la automatización de la aplicación de la misma, reconocida por el autor de GeSCA (Hwang, 2011), siendo el único programa

disponible hasta el momento. Se destaca la utilización de variables manifiestas con escala ordinal, dada su amplia utilización en distintas áreas de la psicología; la incorporación de variables manifiestas formativas considerando su utilización en el ámbito de las ciencias económicas; y la utilización del método de estimación ML para ACOV, el que ha demostrado ser el más eficiente cuando se utiliza un modelo que incluye variables manifiestas formativas (Ringle, Götz, Wetzles, y Wilson, 2009).

En los capítulos teóricos se presentará un diagrama de rutas con la nomenclatura propia de cada técnica de MEE. Sin embargo, por simplicidad y considerando lo virtualmente inevitable de comenzar por definir un modelo de rutas con variables latentes convencional (McDonald, 1996), en la parte práctica de este trabajo se utilizará la nomenclatura de ACOV, tanto en los diagramas de rutas, como en el desarrollo de los dos estudios que incorpora dicho apartado.



## Capítulo 1. Análisis de Rutas (PA)

El origen de los MEE se encuentra en el Análisis de Rutas (Path Analysis, PA), el que constituye una metodología para analizar sistemas de ecuaciones estructurales (Bollen, 1989). Esta metodología fue desarrollada por el biómetra Sewall Wright, quien la plantea por primera vez en 1918, para analizar datos sobre factores que influyen en el crecimiento de conejos, obtenidos por Castle (1914). Wright formaliza su propuesta y presenta nuevas características, como el diagrama de rutas, a partir de 1920 (Wright, 1920, 1921). Ampliando y clarificando su aplicación, a partir de 1932 (Wright, 1932, 1934).

De acuerdo con Bollen (1989), se pueden distinguir tres elementos esenciales del PA:

1. El diagrama de rutas (representación gráfica de un sistema de ecuaciones simultáneas).
2. La descomposición de varianzas, covarianzas y correlaciones en términos de parámetros del modelo.
3. La distinción entre los efectos directos, indirectos y totales entre las variables.

La primera aplicación de este método de análisis en psicología fue realizada por Bárbara Burks (1928), el que cobra relevancia en la estadística, economía y ciencias sociales a partir de 1960, gracias a los trabajos de Blalock (1961, 1962, 1964), Duncan (1966) y Duncan y Hodge (1963). Además, su difusión en las ciencias sociales fue ampliada, a partir de 1970, gracias al trabajo de Goldberger (1971) y Jöreskog (1970a, 1971) en el desarrollo del análisis de estructuras de covarianzas, quienes extienden la aplicación del PA a análisis de modelos con variables latentes o

factores; y los trabajos de Duncan (1975), Jöreskog (1970b), Jöreskog y van Thillo (1972), Keesling (1972), Werst y Linn (1970) y Wiley (1973) en el desarrollo de los MEE, ampliando la aplicación del PA con variables latentes al análisis de las relaciones entre variables latentes.

El PA “se inicia con un diagrama, basado en una teoría, en el que se señalan por medio de flechas las relaciones de influencia entre las variables consideradas” (Lévy y Oubiña, 2006, p.175).

A pesar de lo acertado del planteamiento de Lévy y Oubiña (2006), cabe señalar que la primera utilización del PA realizada por Wright (1918) no incorpora el diagrama de rutas, ilustración de las relaciones entre las variables que es incluida en sus publicaciones posteriores (Wright, 1920, 1921, 1923, 1932 y 1934).

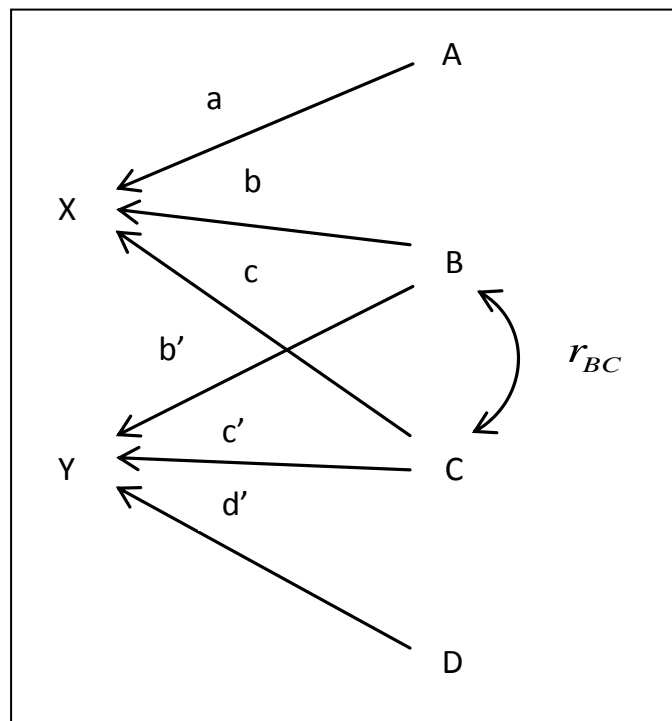


Figura 1. Ejemplo de diagrama ilustrando dos efectos (XY) determinados en parte por las mismas causas correlacionadas (BC) (Wright, 1920).

Los diagramas de rutas representan sistemas de relaciones entre variables. La Figura 1 ilustra un sistema en el cual las variaciones en dos cantidades X e Y son determinadas en parte por causas independientes, como lo son A y D; y en parte por causa comunes, como lo son B y C. Las causas comunes B y C se encuentran correlacionadas ( $r_{BC}$ ). Se asume que todas las relaciones son aproximadamente lineales y que las influencias de las causas están combinadas aproximadamente de forma aditiva (Wright, 1920). X e Y están conectadas por cuatro cadenas independientes, X-B-Y, X-C-Y, X-B-C-Y y X-C-B-Y. Sumando los productos de los coeficientes de rutas (los efectos directos indicados por letras minúsculas), y utilizando  $r_{BC}$  para representar aquellas conectando B y C (Wright, 1923), se obtiene que:

$$r_{XY} = bb' + cc' + br_{BC}c' + cr_{BC}b' \quad (1.1)$$

Esta fórmula puede ser escrita de forma general y conveniente, utilizando los predictores de X, como:

$$r_{XY} = ar_{AY} + br_{BY} + cr_{CY} \quad (1.2)$$

Las relaciones entre las variables, habitualmente incorporadas en un diagrama de rutas (Figura 2) pueden incluir: (1) una relación causal, representada por un sendero directo entre una variable independiente A (exógena) y una variable dependiente X (endógena), (2) una relación de correlación a través de una variable causal común conocida A, (3) una relación de correlación con la totalidad de causas desconocidas, o (4) una relación de causalidad complicada por la correlación de causas comunes desconocidas.

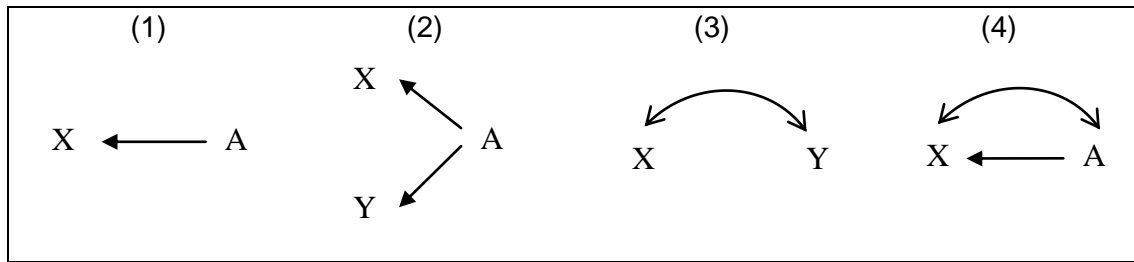


Figura 2. Relaciones entre variables representadas en un diagrama de rutas.

Una de las propiedades de los coeficientes de ruta consiste en que si una variable es totalmente determinada por un grupo de otras variables (Figura 3), la suma de los cuadrados de los coeficientes de rutas dirigidos a ella, más ciertos términos expresando determinación conjunta por las variables correlacionadas, es igual a la unidad. Un término conjunto de este tipo es dos veces el producto de dos coeficientes de rutas multiplicado por el coeficiente de correlación entre las dos variables en cuestión (Wright, 1920, 1923):

$$a^2 + b^2 + c^2 + 2bc r_{BC} = 1 \quad (1.3)$$

Debido a esta propiedad el cuadrado de los coeficientes de ruta entrega una medida del grado de determinación. Cada uno mide la porción de varianza (desviación estándar al cuadrado) explicada por el factor en cuestión (Wright, 1923).

Expresando la correlación de una variable (X) consigo misma, en términos de coeficientes de rutas, de acuerdo a la ecuación (1.1), se obtiene la ecuación (1.3). Además, la ecuación (1.3) puede ser expresada en la siguiente forma simple, la cual puede ser fácilmente generalizada:

$$ar_{AX} + br_{BX} + cr_{CX} = 1 \quad (1.4)$$

La definición de un coeficiente de ruta, incluyendo su relación con la desviación típica y la correlación expresada en las ecuaciones (1.1) y (1.2), con la determinación expresada en las ecuaciones (1.3) y (1.4), es esencialmente la teoría de los coeficientes de rutas (Wright, 1923).

Lévy y Oubiña (2006) señalan que entre los supuestos básicos del PA se incluye que: el modelo debe ser considerado como un sistema cerrado en cuanto contiene explícitamente todas las variables importantes en la determinación causal, y el modelo debe ser recursivo, es decir, que la influencia de una variable sobre otra es asimétrica y, por lo tanto, se ejerce en un solo sentido causal. Además, los autores destacan una serie de supuestos que asumiría el PA, entre los cuales se encuentra que:

- La escala de medición empleada debe ser de intervalo, aunque la introducción de datos ordinales no suele plantear problemas en el análisis.
- Existe una baja multicolinealidad entre las variables.
- Las variables independientes no presentan ningún error de medición.
- Las varianzas de las distribuciones de las variables endógenas del modelo son iguales (supuesto de homocedasticidad).

Los autores plantean además que “el análisis PATH es una aplicación particular del Análisis de Estructuras de Covarianzas que utiliza estrictamente variables observables estableciendo relaciones entre ellas” (Lévy y Oubiña, 2006, p.190). Aplicación a la que Bollen (1989) denomina “Modelos de ecuaciones estructurales con variables observables”.

Un PA, así como las técnicas de MEE desarrolladas posteriormente, se encuentran compuestas como máximo por un conjunto de tres ecuaciones simultáneas, estimadas al mismo tiempo (Golob, 2003):

1. Un modelo de medida (o submodelo) para las variables endógenas (dependientes).
2. Un (sub)modelo de medida para las variables exógenas (independientes).
3. Un (sub)modelo estructural que especifica las relaciones entre las variables endógenas y exógenas, todos los cuales son estimados simultáneamente.

Bollen y Long (1993), argumentan que la mayoría de las aplicaciones de MEE incluyen cinco pasos en su desarrollo:

1. Especificación del modelo. Implica la formulación del modelo basándose en una teoría o en investigaciones previas.
2. Identificación. Implica la determinación de la posibilidad de encontrar valores únicos para los parámetros.
3. Estimación. Implica la selección de la técnica para la estimación de parámetros, de acuerdo con las propiedades de distribución de las variables a analizar.
4. Comprobación del ajuste. Implica la evaluación de la consistencia entre el modelo y los datos.
5. Reespecificación. Implica la modificación en la especificación del modelo.

Cabe señalar que si en el paso 4 se evidencia un buen ajuste, no es necesario continuar con el paso 5. En caso contrario, una vez reespecificado el modelo se deberán realizar nuevamente los pasos a partir del paso número 2.





## Capítulo 2. Análisis de Estructuras de Covarianzas (ACOV)

Karl Jöreskog (1970b), plantea un modelo general para el análisis de estructuras de covarianzas (ACOV), el cual fue implementado, con la colaboración de Gunnar Gruvaeus y Marielle van Thillo (Jöreskog, Gruvaeus y van Thillo, 1970), en el programa computacional denominado ACOVS (Analysis of COVariance Structures). Bajo la asunción de normalidad multivariante de las variables observadas, Jöreskog aplica el método de máxima verosimilitud (maximum likelihood, ML) para la estimación de parámetros en el ACOV. Este modelo general fue ampliado e implementado posteriormente en el programa LISREL (Jöreskog y van Thillo, 1972), el cual es ampliamente utilizado y actualmente continua siendo actualizado de manera periódica (p.e. Jöreskog y Sörbom, 2006).

En el ACOV, la matriz de covarianzas de las variables observadas es modelada como una función de los parámetros de un modelo de ecuaciones estructurales hipotético. Si el modelo es correcto y se conocen los parámetros del modelo estructural, la matriz de covarianzas poblacional podría ser reproducida exactamente a partir de la combinación de los parámetros del modelo (Ruiz, 2000).

El ACOV incorpora determinadas características al diagrama de rutas, tales como: graficar las variables latentes con círculos u óvalos y las variables manifiestas con cuadrados o rectángulos. Además incorpora tres tipos de relaciones entre variables (Figura 3).

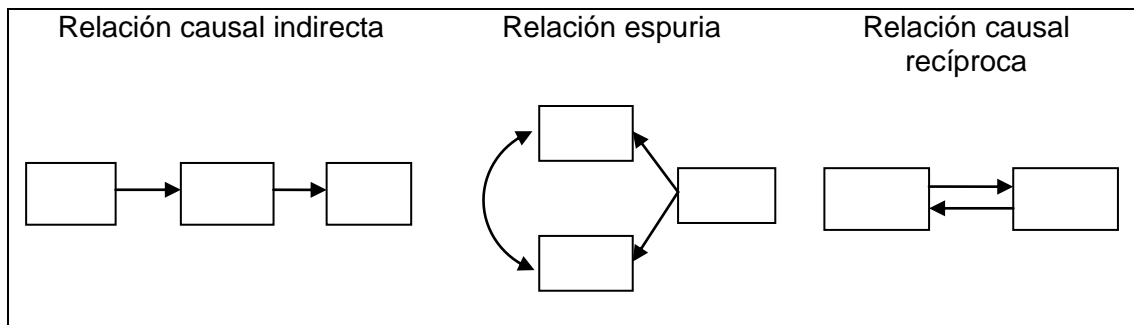


Figura 3. Relaciones entre variables incorporadas por el ACOV.

El método general para analizar datos de acuerdo con un modelo general para estructuras de covarianza planteado por Jöreskog (1970b, 1971, 1979a), para el análisis factorial, asume que la matriz poblacional de varianzas-covarianzas  $\Sigma(p \times p)$  =  $(\sigma_{ij})$  de un conjunto de variables tiene la forma:

$$\Sigma = \mathbf{B}(\mathbf{A}\Phi\mathbf{A}' + \Psi^2)\mathbf{B}' + \Theta^2 \quad (2.1)$$

donde  $\mathbf{B}$  es una matriz  $p \times q$  que contiene los efectos ( $\beta$ ) de las  $q$  variables latentes sobre las  $p$  variables manifiestas,  $\mathbf{A}$  es una matriz de orden  $q \times r$  que contiene los coeficientes de regresión ( $\lambda$ ) de las  $q$  variables latentes endógenas sobre las  $r$  variables latentes exógenas, la matriz simétrica  $\Phi$  de dimensiones  $r \times r$  contiene las correlaciones ( $\phi$ ) entre las variables latentes exógenas, y las matrices diagonales  $\Psi$  de orden  $(q \times q)$  que contiene las correlaciones entre los errores de las variables latentes endógenas ( $\psi$ ) y  $\Theta$  de dimensiones  $(p \times p)$  que contiene los errores de medida de las variables manifiestas ( $\theta$ ), son matrices de parámetros.

Los parámetros pueden ser de tres tipos: (i) parámetros fijos, a los cuales se les ha asignado un valor determinado, (ii) parámetros restringidos, los cuales son

desconocidos, pero se ha determinado que son iguales a uno o más de los otros parámetros, y (iii) parámetros libres, los cuales son desconocidos y no son restringidos a ser iguales a ningún otro parámetro (Batista-Foguet y Coenders, 2000; Jöreskog, 1970b, 1971, 1979a, 1979b; Jöreskog y Goldberger, 1975; Jöreskog y van Thillo, 1972).

*“En cualquier aplicación del modelo de la ecuación (2.1) el número de variables  $p$  esta dado por los datos,  $q$  y  $r$  están dados por la aplicación particular que el investigador tiene en mente”* (Jöreskog, 1971, p. 110; 1979a, p. 47).

La estructura de la matriz de covarianza surge cuando las variables observadas  $\mathbf{x}(p \times 1)$  son de la forma:

$$\mathbf{x} = \boldsymbol{\mu} + \mathbf{B}\boldsymbol{\Lambda}\boldsymbol{\xi} + \mathbf{B}\boldsymbol{\zeta} + \mathbf{e} \quad (2.2)$$

donde  $\boldsymbol{\xi}(r \times 1)$ ,  $\boldsymbol{\zeta}(q \times 1)$ , y  $\mathbf{e}(p \times 1)$  son vectores aleatorios latentes no correlacionados, en general no observados, con vectores de media cero y matrices de dispersión  $\boldsymbol{\Phi}$ ,  $\boldsymbol{\Psi}^2$ , y  $\boldsymbol{\Theta}^2$ , respectivamente, y donde  $\boldsymbol{\mu}$  es el vector de medias de  $\mathbf{x}$  (Jöreskog, 1979a).

El modelo general es ampliado para estimar coeficientes desconocidos en una serie de ecuaciones estructurales lineales (Jöreskog, 1970a, 1979b), realizando algunos cambios de notación.

El modelo de variables latentes (modelo estructural) contiene las ecuaciones correspondientes a las relaciones entre las variables no observables:

$$\boldsymbol{\eta} = \boldsymbol{\beta}\boldsymbol{\eta} + \boldsymbol{\Gamma}\boldsymbol{\xi} + \boldsymbol{\zeta} \quad (2.3)$$

donde  $\boldsymbol{\eta}$  es un vector  $m \times 1$  que contiene las variables latentes endógenas,  $\boldsymbol{\xi}$  es un vector  $n \times 1$  que contiene las variables latentes exógenas,  $\boldsymbol{\beta}$  es una matriz  $m \times m$  que contiene los coeficientes de regresión de las variables endógenas sobre las variables endógenas,  $\boldsymbol{\Gamma}$  es una matriz  $m \times n$  que contiene los coeficientes de regresión de las variables endógenas sobre las variables exógenas, y  $\boldsymbol{\zeta}$  es un vector  $m \times 1$  que contiene los errores de regresión que resultan de predecir las variables endógenas a partir de las exógenas. Se asume que las variables latentes contenidas en  $\boldsymbol{\eta}$  y  $\boldsymbol{\xi}$  no están correlacionadas y que la matriz  $\boldsymbol{\beta}$  no es singular (Bollen, 1989; Jöreskog, 1979b; Kaplan, 2009; Ruiz, 2000).

El modelo de medida incluirá tantas ecuaciones como variables observadas existan. Las ecuaciones habitualmente son agrupadas en variables endógenas ( $\mathbf{y}$ ) y variables exógenas ( $\mathbf{x}$ ):

$$\mathbf{y} = \boldsymbol{\Lambda}_y \boldsymbol{\eta} + \boldsymbol{\varepsilon} \quad (2.4)$$

donde  $\mathbf{y}$  es un vector  $p \times 1$  que contiene las variables aleatorias correspondientes a las variables observables,  $\boldsymbol{\Lambda}_y$  es una matriz de orden  $p \times m$  que contiene las saturaciones de las variables manifiestas en las variables latentes dependientes, y el vector  $\boldsymbol{\varepsilon}$ , de dimensiones  $p \times 1$ , contiene las variables aleatorias correspondientes a los errores de medida (Bollen, 1989; Jöreskog, 1979b; Kaplan, 2009; Ruiz, 2000).

$$\mathbf{x} = \boldsymbol{\Lambda}_x \boldsymbol{\xi} + \boldsymbol{\delta} \quad (2.5)$$

donde  $\mathbf{x}$  es un vector  $q \times 1$  que contiene las variables aleatorias correspondientes a las variables observables,  $\Lambda_x$  es una matriz de orden  $q \times n$  que contiene las saturaciones de las variables observables en las variables latentes independientes, y el vector  $\delta$ , de dimensiones  $q \times 1$ , contiene las variables aleatorias correspondientes a los errores de medida.

Jamshidian y Mata (2007), señalan que los modelos de ecuaciones estructurales principalmente consisten en situar estructuras sobre la matriz de covarianzas de la población  $\Sigma$ . Estas estructuras resultan de una combinación de modelos de medida y modelos de variables latentes.

El planteamiento de Jamshidian y Mata (2007) sobre los MEE, concuerda con la ecuación que Bollen (1989) denomina hipótesis fundamental:

$$\Sigma = \Sigma(\theta) \quad (2.6)$$

donde  $\Sigma$  es la matriz de varianzas y covarianzas de las variables observadas de la población,  $\theta$  es un vector que contiene los parámetros del modelo, y  $\Sigma(\theta)$  es la matriz de varianzas y covarianzas escrita como función de  $\theta$ .

$$\Sigma(\theta) = \begin{bmatrix} \Sigma_{yy}(\theta) & \Sigma_{yx}(\theta) \\ \Sigma_{xy}(\theta) & \Sigma_{xx}(\theta) \end{bmatrix} \quad (2.7)$$

$$= \begin{bmatrix} \Lambda_y [(\mathbf{I} - \mathbf{B})^{-1} (\Gamma \Phi \Gamma' + \Psi)] [(\mathbf{I} - \mathbf{B})^{-1}]' \Lambda_y' + \Theta_\varepsilon & \Lambda_y (\mathbf{I} - \mathbf{B})^{-1} \Gamma \Phi \Lambda_x' \\ \Lambda_x \Phi \Gamma' [(\mathbf{I} - \mathbf{B})^{-1}]' \Lambda_y' & \Lambda_x \Phi \Lambda_x' + \Theta_\delta \end{bmatrix}$$

donde  $\mathbf{I}$  es una matriz identidad de dimensiones  $m \times m$ ;  $\mathbf{B}$  es una matriz  $m \times m$  que contiene los efectos ( $\beta$ ) de variables endógenas sobre otras variables endógenas,  $\Phi$  es una matriz simétrica, de dimensiones  $n \times n$ , que contiene las correlaciones ( $\phi$ ) entre las variables latentes exógenas,  $\Psi$  es una matriz diagonal de orden ( $m \times m$ ) que contiene las correlaciones entre los errores de las variables endógenas latentes,  $\Theta_\varepsilon$  es una matriz de dimensiones ( $p \times p$ ) que contiene los errores de medida de las variables manifiestas endógenas, y  $\Theta_\delta$  es una matriz de dimensiones ( $q \times q$ ) que contiene los errores de medida de las variables manifiestas exógenas

La matriz  $\Sigma(\theta)$ , expresada en la ecuación (2.7), es la matriz de covarianzas reproducida en el modelo.

En la práctica, para contrastar la hipótesis fundamental es necesario comenzar por la *Especificación* del Modelo, estableciendo las variables que lo conforman y el tipo de relación existente entre ellas a partir de la teoría, definiendo así el modelo de medida y el modelo estructural, determinando además los tipos de parámetros que serán incluidos (libres, restringidos o fijos). Al respecto Mulaik (2009) precisa que es un error hablar de “modelo mal especificado” cuando se libera un parámetro que es cero en la población, ya que al liberar un parámetro nada es especificado, de hecho, menos es especificado del modelo. Por lo que en la presente tesis doctoral se considerarán los niveles de especificación de modelos: Más especificado y Menos especificado.

En cuanto al tamaño muestral adecuado para la aplicación de ACOV, existen distintos criterios. Así, por ejemplo Anderson y Gerbing (1988) plantean la necesidad de contar con un tamaño muestral igual o superior a 150. Iacobucci (2010) señala que

a pesar que los MEE pueden tener un buen desempeño con tamaños muestrales pequeños (p.e. 50 o 100), la regla habitualmente utilizada de un tamaño superior a 200 (Marsh, Hau, Balla y Grayson, 1998) puede resultar conservadora y simplista, señalando la conveniencia de considerar el índice de ajuste, la cantidad de variables y la relación entre variables para determinar el tamaño muestral adecuado (Kim, 2005). Una regla habitualmente utilizada y aceptada es utilizar 10 sujetos por cada indicador considerado en el modelo (Browne y Cudeck, 1993; Gerbing y Anderson, 1985)

Después de la especificación, se debe comprobar la *Identificabilidad* del modelo, determinando si las covarianzas entre las variables observables facilitan información suficiente para estimar unívocamente los parámetros del modelo. La estimación se encuentra relacionada con la resolución del sistema expresando los parámetros (incógnitas) en función de las varianzas y covarianzas (ecuaciones). Así la diferencia entre el número de varianzas y covarianzas y el de parámetros a estimar determinan los ( $g$ ) grados de libertad del modelo (Batista-Foguet y Coenders, 2000).

De acuerdo, Batista-Foguet y Coenders (2000), según el valor de  $g$ , los modelos estructurales pueden clasificarse en:

- 1) Nunca identificados ( $g < 0$ ): Modelos en los que los parámetros podrían tomar infinitos valores, por lo que se encuentran indeterminados.
  
- 2) Posiblemente identificados ( $g = 0$ ): Modelos en los que puede existir una única solución para los parámetros que iguale la matriz de covarianzas observada e reproducida. Estos modelos son estimables pero su bondad de ajuste será desconocida.

- 3) Posiblemente sobreidentificados ( $g > 0$ ): Modelos que incluyen menos parámetros que varianzas y covarianzas. Sólo estos modelos pueden ser contrastados a partir de los datos.

No existe un conjunto simple de condiciones necesarias y suficientes que permitan evaluar la identificabilidad de cualquier modelo. Una condición necesaria, pero no suficiente, para que un modelo esté identificado es que los grados de libertad sean mayores o iguales a 0. Una serie de condiciones suficientes para una serie de modelos bastante habituales serían:

- Que el modelo sea recursivo (que no incluya una relación causal recíproca), que no incluya efectos entre variables endógenas entre sí, aunque los términos de error de las variables endógenas estén relacionados.
- Que cada variable latente tenga al menos dos variables manifiestas puras (que se relacionan con una única variable latente y cuyos errores de medida son independientes de los errores de otra variable manifiesta).
- Que cada variable latente tenga al menos una variable manifiesta cuya saturación está restringida a un valor distinto de cero (generalmente a la unidad), para fijar la escala de la variable latente.

Los modelos de ecuaciones estructurales asumen tres supuestos estadísticos:

1. Linealidad: Existencia de relaciones lineales entre variables manifiestas y factores; y entre factores. Pudiendo expresar las relaciones mediante un sistema de ecuaciones simultáneas.



2. Normalidad Multivariante: Se asume que la distribución de las fuentes de variación (factores exógenos, términos de error y errores de medida) es normal multivariante.
3. Aislamiento: Ausencia de covariación entre los factores y los términos de error. Además de ausencia de correlación entre los términos de error de predicción y los errores de medida.

Posterior a la comprobación de la identificabilidad y de los supuestos, se encuentra la etapa de la *Estimación*, la cual consiste en la aplicación de una función de ajuste para obtener los valores de los parámetros que generen la mayor proximidad a la equivalencia entre la matriz de covarianzas reproducida de parámetros estructurales ( $\Sigma(\theta)$ ) y la matriz de covarianzas observada ( $S$ ). La función de ajuste se expresa de manera genérica como  $F(S, \Sigma(\theta))$  y tiene las siguientes propiedades:

1. El resultado de  $F(S, \Sigma(\theta))$  debe ser un escalar.
2.  $F(S, \Sigma(\theta)) \geq 0$ .
3.  $F(S, \Sigma(\theta)) = 0$ , sólo si  $\Sigma(\theta) = S$ .
4.  $F(S, \Sigma(\theta))$  es una función continua en  $S$  y  $\Sigma(\theta)$ .

Bollen (1989), en concordancia con Loehlin (1987), plantea 3 funciones de ajuste apropiadas para el modelo general de ecuaciones estructurales, a saber:

1. Función de Máxima Verosimilitud (*Maximum Likelihood*, ML):

$$F_{ML} = \log |\Sigma(\theta)| + \text{tr} \{S \Sigma^{-1}(\theta)\} - \log |S| - (p + q) \quad (2.8)$$

donde  $p$  es el número de variables endógenas observables y  $q$  es el número de variables exógenas observables. Se asume que  $\mathbf{S}$  y  $\Sigma(\boldsymbol{\theta})$  son definidas positivas para evitar que sus determinantes sean nulos y que no existan sus logaritmos. Esta función es asintóticamente insesgada y eficiente, en la medida que se cuenta con muestras grandes que cumplen con la distribución normal multivariante. Además presenta invarianza de escala, por lo que los valores de la función son los mismos al utilizar la matriz de correlaciones y la matriz de covarianzas.

2. Función de Mínimos Cuadrados Generalizados (*Generalized Least Squares*, GLS):

$$F_{GLS} = \left(\frac{1}{2}\right) tr \left\{ \left[ \mathbf{I} - \Sigma(\boldsymbol{\theta}) \mathbf{S}^{-1} \right]^2 \right\} \quad (2.9)$$

Esta función es invariante respecto a la escala, es asintóticamente eficiente y da lugar a estimaciones consistentes de distribución normal multivariante. Sin embargo, cuando la distribución de las variables manifiestas tiene colas demasiado gruesas o delgadas los errores típicos asintóticos y las pruebas de chi cuadrado no son precisos (Bollen, 1989).

3. Función de Mínimos Cuadrados No Ponderados (*Unweighted Least Squares*, ULS):

$$F_{ULS} = \left(\frac{1}{2}\right) tr \left\{ \left[ \mathbf{S} - \Sigma(\boldsymbol{\theta}) \right]^2 \right\} \quad (2.10)$$

Esta función no requiere que las variables presenten una distribución normal multivariante. Sin embargo, es menos eficiente que  $F_{ML}$  y no es invariante respecto a la escala de medida (Bollen 1989).

Por su parte, Jöreskog y Sörbom (1996-2001) plantean dos funciones de ajuste, además de las ya mencionadas, a saber:

4. Mínimos Cuadrados Ponderados (*Weighted Least Squares*, WLS):

$$F_{WLS} = [\mathbf{s} - \boldsymbol{\sigma}(\boldsymbol{\theta})]' \mathbf{W}^{-1} [\mathbf{s} - \boldsymbol{\sigma}(\boldsymbol{\theta})] \quad (2.11)$$

donde  $\mathbf{s}$  es un vector de estadísticos muestrales (correlaciones policóricas),  $\boldsymbol{\sigma}(\boldsymbol{\theta})$  es el vector de elementos de la población del modelo implícito en  $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ ; y  $\mathbf{W}$  es una matriz de pesos definida positiva (Flora y Curran, 2004). Esta función no requiere que las variables presenten una distribución normal multivariante, por lo que puede ser utilizada si algunas variables manifiestas son ordinales o si el modelo incluye variables dicotómicas. Corresponde al método de estimación denominado Asintóticamente Libre de Distribución (*Asymptotically Distribution Free*, ADF) por Browne (1984).

5. Mínimos Cuadrados Ponderados Diagonalizados (*Diagonally Weighted Least Squares*, DWLS): Esta función implica que los parámetros estimados son obtenidos sustituyendo la matriz  $\mathbf{W}$ , por la matriz diagonal  $\mathbf{V}$  en la Ecuación (2.11). Los elementos de esta matriz son las varianzas asintóticas de los coeficientes estimados. Jöreskog y Sörbom (1996-2001) plantean que esta función de ajuste puede ser utilizada cuando se analizan

matrices de correlación. Así, Flora y Curran (2004) sugieren una variación utilizando las estimaciones de las correlaciones policóricas, a la que denominan Mínimos Cuadrados Ponderados Robustos, la cual es más robusta frente a muestras pequeñas.

$$F_{DWLS} = [\mathbf{s} - \boldsymbol{\sigma}(\boldsymbol{\theta})]' \mathbf{V}^{-1} [\mathbf{s} - \boldsymbol{\sigma}(\boldsymbol{\theta})] \quad (2.12)$$

donde  $\mathbf{V} = \text{diag}(\mathbf{W})$ ;  $\mathbf{s}$  es un vector de estadísticos muestrales (correlaciones policóricas),  $\boldsymbol{\sigma}(\boldsymbol{\theta})$  es el vector de elementos de la población del modelo reproducido en  $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ .

Es conveniente considerar que en condiciones de escala de medida discreta y ordinal, como cuando se emplea en los cuestionarios la modalidad de respuesta tipo Likert (1= “completamente de acuerdo”, 2= “de acuerdo”, 3= “ni de acuerdo ni en desacuerdo”, 4= “en desacuerdo”, 5= completamente en desacuerdo”), las covarianzas no constituyen una medida de asociación apropiada, por lo que como punto de partida de la estimación deben emplearse las correlaciones policóricas (Flora y Curran, 2004; Olsson, 1979).

Una vez que se ha obtenido una estimación de los parámetros del modelo se deberán efectuar diagnósticos sobre la adecuación de la solución (signo y significación de parámetros estimados y presencia de casos Heywood) y la bondad de ajuste del modelo (estadísticos globales que contrastan la capacidad del modelo estimado para reproducir la matriz de parámetros muestrales; Ruiz, 2000).

Lévy, Martín y Román (2006) plantean que la evaluación del ajuste del modelo debe realizarse a tres niveles: evaluación del modelo global, evaluación del modelo de medida y evaluación del modelo estructural:

- a) Para la evaluación del modelo global los autores presentan tres medidas de ajuste global complementarias: (1) Ajuste absoluto. Determinan el grado de equivalencia entre la matriz de covarianzas reproducida de parámetros estructurales ( $\Sigma(\theta)$ ) y la matriz de covarianzas observada ( $S$ ), (2) Ajuste incremental. Comparan el ajuste del modelo analizado con un modelo nulo (sin relación entre variables) o con un modelo independiente (modelo base), y (3) Ajuste parsimonioso. Relacionan la bondad de ajuste del modelo con el número de coeficientes estimados necesarios para conseguir dicho nivel de ajuste. Para una revisión de índices de ajustes se puede consultar a Browne y Cudeck (1993), Hu y Bentler (1999), Kaplan (2009), Marsh, Hau y Grayson (2005), Schermelleh-Engel, Moosbrugger y Müller (2003) y Steiger (2007).
  
- b) Para la evaluación del modelo de medida, los autores sugieren examinar 4 propiedades: (1) La validez convergente de los factores latentes, a través de comprobar la presencia de cargas factoriales estandarizadas significativas y superiores a 0,6; (2) La unidimensionalidad de cada factor latente, a través de comprobar el ajuste por separado de cada factor con sus variables manifiestas, y a través de comprobar la validez discriminante entre factores, esta última puede ser analizada a través de tres procedimientos: la revisión de las estimaciones de las correlaciones entre factores, las cuales deben ser inferiores a 0,5; la comparación del estadístico chi-cuadrado del modelo con todas las variables latentes covariando y del mismo modelo pero fijando a 1 la covariación entre las variables latentes que cuya validez discriminante plantee dudas, la validez discriminante será demostrada si chi-cuadrado del

primer modelo es estadísticamente menor; y por último, aplicar el test de intervalos de confianza de  $\pm 2$  errores estándar entre la covarianza de los factores latentes cuya validez discriminante suscita dudas, los intervalos de confianza que no incluyan la unidad serán indicativos de la existencia de validez discriminante; y (3) La fiabilidad de los constructos, a través de examinar la varianza extraída, que refleja la cantidad total de la varianza de las variables manifiestas tenida en cuenta por el constructo latente, la cual debe ser superior a 0,5; y (4). La fiabilidad de las variables manifiestas, a través de examinar el coeficiente de determinación o correlación múltiple al cuadrado, la cual mide el porcentaje de varianza común entre la variable observable y la variable latente, debiendo ser superior a 0,5 (Lévy, Martín y Román, 2006).

- c) Para la evaluación del ajuste del modelo estructural, los autores proponen identificar si todos los estimadores del modelo son significativos. En caso contrario, suprimir la relación entre variables latentes, siempre y cuando ello no distorsione la teoría en que se basa el modelo (Lévy, Martín y Román, 2006).

Finalmente, una vez evaluado el ajuste del modelo, considerando la posibilidad de estar en presencia de un modelo con mal ajuste global o la presencia de evaluaciones del modelo de medida o del modelo estructural que sugieran agregar o quitar vínculos entre variables, se podrá realizar la reespecificación del modelo, siempre y cuando las modificaciones a realizar fuesen avaladas por la teoría. En este ámbito, existen indicadores empíricos que pueden orientar las modificaciones necesarias para mejorar el ajuste del modelo global:

- Residuos estandarizados: Un porcentaje de residuos estandarizados estadísticamente significativos superior al 5% indicará diferencias entre la matriz de covarianzas estimada y la matriz de covarianzas observada. Los parámetros o variables con mayores valores de residuos estandarizados asociados podrán ser modificados o eliminados (Lévy, Martín y Román, 2006).
- Índices de modificación: Expresan la reducción que experimentaría Chi-cuadrado en el caso de incorporar una nueva relación en el modelo. Valores superiores a 3,84 sugerirán la inclusión de una relación entre variables (Lévy, Martín y Román, 2006).

Además, para determinar la necesidad de liberar o restringir parámetros del modelo (reespecificar el modelo) se puede utilizar el test de multiplicadores de Lagrange y el test Wald, respectivamente o la razón de verosimilitud (Bollen, 1989; Engle, 1984; García Ferrer, 1984, Levy, Martín y Román, 2006). Bollen (1989) señala que los índices de modificación propuestos por Jöreskog y Sörbom son equivalentes a los multiplicadores de Lagrange.

Al realizar una reespecificación del modelo se debe continuar con la comprobación de la identificabilidad, supuestos estadísticos, estimación de parámetros, comprobación de la adecuación de la solución y la bondad de ajuste del modelo. Sería esperable que el modelo reespecificado presente un mejor ajuste que el modelo sin reespecificar; independientemente del cumplimiento de lo señalado. El proceso finalizaría con el informe de los resultados encontrados, para la redacción del cual se sugiere utilizar las directrices planteadas por Hoyle y Panter (1995), Schreiber, Stage, King, Nora y Barlow (2006) y Steiger (1988).

Temas avanzados en el ámbito de los MEE tales como, análisis de invarianza factorial multigrupo, MEE multinivel, modelos longitudinales, comparación de medias del modelo estructural y modelos de curva latente pueden ser encontrados en las publicaciones de Hancock y Mueller (2006), Kaplan (2009), Kline (2011), Lévy y Varela (2006) y Mulaik (2009), entre otros.



## Capítulo 3. Modelo de Rutas mediante Mínimos Cuadrados Parciales (PLS-PM)

Herman Wold comienza a desarrollar la idea de Mínimos Cuadrados Parciales (Partial Least Squares, PLS) a partir de 1954, cuando plantea una generalización de los sistemas recursivos (Wold, 1954). Formaliza su propuesta al introducir el algoritmo NILES (Nonlinear Iterative LEeast Squares) para calcular componentes principales (Wold, 1966), a través de una secuencia iterativa de regresiones de mínimos cuadrados ordinarios (OLS). Algoritmo que, posteriormente, extiende para aplicarlo a más de tres bloques de variables latentes; y denomina Mínimos Cuadrados Parciales Iterativos No Lineales (Non-linear Iterative PArtial Least Squares, NIPALS; Wold, 1973, 1974, 1975a, 1975b; Wold y Lyttkens, 1969). Posteriormente, Wold (1977, 1980) se centra en la construcción de modelos, abandonando la denominación NIPALS y variando entre la denominación PLS (Wold, 1980) y soft modeling (Wold, 1982a) optando finalmente por la primera (Wold, 1982b, 1985), la cual es actualmente denominada Modelo de Rutas mediante Mínimos Cuadrados Parciales (Partial Least Squares-Path Modeling, PLS-PM; Esposito Vinzi, Trinchera y Amato, 2010; Henseler, Ringle y Sinkovics, 2009; Tenenhaus, Esposito Vinzi, Chatelin y Lauro, 2005). Cabe destacar el aporte de Jan-Bernd Lohmöller al desarrollo de PLS-PM, quien propuso e implementó varias extensiones a la propuesta original de Wold, en cuanto a la estandarización de las variables manifiestas y las variables latentes (Lohmöller, 1987, 1989).

Algunas de las bondades de la propuesta de Wold y que sustentan la denominación de modelo blando, se encuentran en la posibilidad de obtener soluciones fiables con menos restricciones que ACOV (con el método de estimación de parámetros ML) en cuanto a la distribución de los datos y el tamaño de la muestra. Lo que suele favorecer el considerar a PLS-PM como una alternativa para ser utilizada

en estudios aplicados en ciencias sociales, del comportamiento y de la salud, los cuales suelen incumplir algunos de los requisitos de ACOV. Sin embargo, Marcoulides (2006), sugiere evitar el error de la creencia que PLS-PM debe ser utilizado en todos los casos en que el tamaño muestral es pequeño, argumentando que las estimaciones de PLS-PM tenderán a valores verdaderos en la medida que el tamaño muestral y el número de indicadores tiendan al infinito (McDonald, 1996; Hui y Wold, 1982). Esta sugerencia, aunque conveniente de ser tomada en cuenta a la hora de elegir entre las técnicas de MEE, no incluye la necesidad de considerar finalidad de la técnica de análisis, la que debe ser coherente con la finalidad de la investigación que esté llevando a cabo. Así, un aspecto importante a considerar es que PLS-PM se centra en maximizar la varianza de las variables dependientes explicada por las variables independientes, en vez de reproducir la matriz de covarianzas empírica (Haenlein y Kaplan, 2004), como en el caso del ACOV. Por lo que PLS-PM debe ser elegido cuando objetivo de la investigación esté centrado en la predicción de las variables dependientes.

Un modelo PLS-PM se encuentra constituido por dos (sub)modelos, el Modelo Interno (Modelo Estructural) que vincula las variables latentes entre sí; y el Modelo Externo (Modelo de Medida) que vincula las variables manifiestas con las variables latentes (Figura 4).

Esposito Vinzi, Trinchera y Amato (2010) plantean que en el marco del PLS-PM el modelo estructural puede ser escrito como:

$$\xi_j = \beta_{0j} + \sum_{q: \xi_q \rightarrow \xi_j} \beta_{qj} \xi_q + \zeta_j \quad (3.1)$$

donde  $\xi_j$  ( $j=1, \dots, J$ ) es la variable latente endógena genérica,  $\beta_{qj}$  es el coeficiente de la ruta genérica interrelacionando la  $q$ -ésima variable latente exógena a la  $j$ -ésima variable latente endógena, y  $\zeta_j$  es el error en la relación interna (estructural).

La formulación del Modelo de Medida, depende de la dirección de la relación entre las variables latentes y las correspondientes variables manifiestas (Fornell y Bookstein, 1982). La cual puede definir tres tipos de modelos de medida: modelo reflectivo, modelo formativo o modelo mixto.

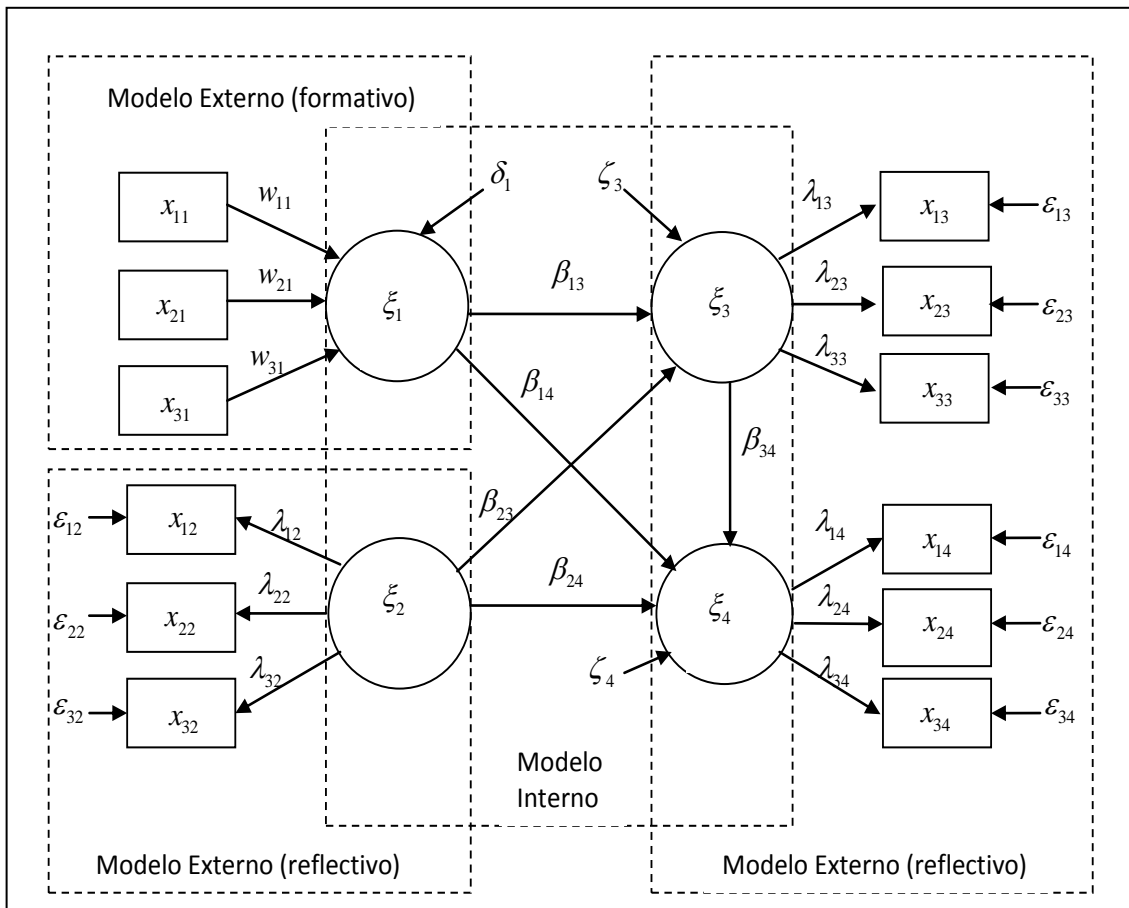


Figura 4. Ejemplo de Diagrama de rutas de un Modelo de Rutas mediante Mínimos Cuadrados Parciales.

En un Modelo de Medida *Reflectivo*, el conjunto de variables manifiestas relacionado con una variable latente, se asume que mide un único concepto

subyacente. Cada variable manifiesta refleja (es un efecto de) la variable latente correspondiente, y juega un rol de variable endógena en el conjunto específico del Modelo de Medida.

En el Modelo de Medida Reflectivo las variables manifiestas relacionadas a una misma variable latente deberían covariar, es decir, cambios en una variable manifiesta implican cambios en las otras. Cada variable manifiesta está relacionada con la variable latente correspondiente a través de un modelo de regresión simple:

$$\mathbf{x}_{pq} = \lambda_{p0} + \lambda_{pq}\xi_q + \varepsilon_{pq} \quad (3.2)$$

donde  $\lambda_{pq}$  es la carga asociada a la  $p$ -ésima variable manifiesta en la  $q$ -ésima variable latente, y el término de error  $\varepsilon_{pq}$  representa la imprecisión en el proceso de medida.

Sobre el modelo de la ecuación (3.2) se aplica la hipótesis de la *condición de especificación del predictor*:

$$E(\mathbf{x}_{pq} / \xi_q) = \lambda_{p0} + \lambda_{pq}\xi_q \quad (3.3)$$

Esta hipótesis implica que el residual  $\varepsilon_{pq}$  tiene media igual a 0 y es independiente de (está incorrelacionado con) la variable latente  $\xi_q$ .

De acuerdo con lo señalado por Tenenhaus, Esposito Vinzi, Chatelin y Lauro (2005), debido a que un grupo reflectivo de variables, refleja un (único) constructo latente, debería ser homogéneo y unidimensional. Los autores plantean tres herramientas disponibles para valorar dichos aspectos: alpha de Cronbach (superior a

0,7), rho de Dillon-Goldstein (superior a 0,7) y análisis de componentes principales del grupo, utilizando la regla de Kaiser (primer autovalor superior a 1 y los demás inferiores a 1). Por su parte, McDonald (1981) refiere que la proporción de varianza debido al primer componente principal o al primer factor común es un criterio habitual e insatisfactorio para la unidimensionalidad. Aunque para probar la unidimensionalidad existen una amplia variedad de procedimientos - Hattie (1984) identifica 87 índices y Martínez Arias (1995) describe 13 de ellos - suele utilizarse el análisis factorial (Muñiz y Cuesta, 1993) de tipo confirmatorio probando el ajuste de un modelo unidimensional y evaluando la matriz de residuos e índices de bondad ajuste, o de tipo exploratorio con análisis paralelo, evaluando la cantidad de factores retenidos (Abad, Olea, Ponsoda y García, 2011; Ferrando, 1996).

En el Modelo de Medida *Formativo*, cada variable o cada subgrupo de variables manifiestas representa una dimensión diferente del concepto subyacente. A diferencia del modelo reflectivo, el modelo formativo no asume homogeneidad ni unidimensionalidad del grupo de variables y las variables manifiestas no requieren covariar, aunque suelen encontrarse correlacionadas. La variable latente es definida como una combinación lineal de las variables manifiestas correspondientes, así cada variable manifiesta es una variable exógena en el Modelo de Medida. Este modelo de medida puede ser expresado como:

$$\xi_q = \sum_{p=1}^{P_q} \omega_{pq} x_{pq} + \delta_q \quad (3.4)$$

donde  $\omega_{pq}$  es el coeficiente que une cada variable manifiesta a la variable latente correspondiente (peso de relación), y el término de error  $\delta_q$  representa la fracción de

la variable latente correspondiente no considerada por el grupo de variables manifiestas.

Se aplica la *condición de especificación del predictor* sobre la ecuación (3.4):

$$E(\xi_q | x_1, \dots, x_{p_q}) = \sum_{p=1}^{P_q} \omega_{pq} x_{pq} \quad (3.5)$$

Esta hipótesis implica que el residual  $\delta_q$  tiene media igual a 0 y es independiente de (está incorrelacionado con) la variable manifiesta  $x_{pq}$ .

Finalmente, el modelo mixto, es una mezcla de ambos modelos (reflectivo y formativo) en el mismo grupo de variables manifiestas.

Independientemente del tipo de modelo de medida, las puntuaciones estandarizadas de las variables latentes ( $\xi_q$ ) asociados a la q-ésima variable latente ( $\xi_q$ ) son calculados, hasta lograr la convergencia, como una combinación lineal de su bloque de variables manifiestas:

$$\hat{\xi}_q = \sum_{p=1}^{P_q} w_{pq} x_{pq} \quad (3.6)$$

donde las variables  $x_{pq}$  están centradas y  $w_{pq}$  son los pesos externos, que son generados hasta lograr la convergencia del algoritmo, detallado más adelante, y luego transformados para producir puntuaciones estandarizadas de variables latentes.

Es importante no confundir los pesos de relación  $\omega_{pq}$  de la ecuación (3.4) y (3.5), con los pesos externos  $w_{pq}$  de la ecuación (3.6). Los primeros implican que en PLS-PM cualquier variable latente es definida como una suma ponderada de sus variables manifiestas, sin afectar la dirección de la relación entre la variable latente y sus variables manifiestas. Dicha dirección determina la forma en que serán estimados los pesos usados en la ecuación 3.6 en el algoritmo.

“En PLS-PM, un procedimiento iterativo permite estimar los pesos externos ( $w_{pq}$ ) y las puntuaciones de las variables latentes ( $\xi_q$ ). El procedimiento de estimación es llamado parcial, debido a que considera un conjunto de variables cada vez, por medio de regresiones lineales simples y múltiples alterantes. Los coeficientes de rutas ( $\beta_{qj}$ ) son estimados posteriormente por medio de regresiones entre las puntuaciones estimadas de las variables latentes en conformidad con la red de relaciones estructurales especificada” (Esposito Vinci, Trinchera y Amato, 2010, pp. 52).

La estimación de los pesos externos es llevada a cabo a través de una alternancia entre los pasos de estimación externos e internos, a través del algoritmo de punto fijo (FP; Lyttkens, 1968, 1973; Wold, 1965, 1981); en el cual un conjunto de parámetros del modelo es dividido en subconjuntos y cada subconjunto es parcialmente estimado a través de mínimos cuadrados ordinarios (OLS) con otros subconjuntos fijos. Esta estimación OLS es iterativa hasta lograr la convergencia.

El algoritmo de PLS-PM, en concordancia con la propuesta original de Wold (1980, 1982a, 1985), consta de 3 etapas. En las 2 primeras el algoritmo trabaja con variables estandarizadas (normalizadas), con media igual a 0 y desviación estándar

igual a 1. La primera etapa es un procedimiento que busca obtener los pesos externos. La segunda busca obtener los coeficientes de carga; y la última, la obtención de los coeficientes de rutas:

a. Etapa 1. Estimación iterativa de pesos y puntuaciones de variables latentes:

Esta etapa consta de cuatro pasos. El procedimiento trabaja sobre las variables manifiestas centradas (o estandarizadas) y comienza determinando pesos iniciales  $w_{pq}$  arbitrariamente (por ejemplo, utilizar todos los pesos con valor igual a 1).

El *Paso 1*, consiste en la estimación externa, en la cual cada variable latente es estimada, como una combinación lineal de sus variables manifiestas:

$$\mathbf{v}_q \propto \pm \sum_{p=1}^{P_q} w_{pq} \mathbf{x}_{pq} = \pm \mathbf{X}_q \mathbf{w}_q \quad (3.7)$$

donde  $\mathbf{v}_q$  es la estimación externa estandarizada de la  $q$ -ésima variable latente  $\xi_q$ ,  $\propto$  implica que el lado izquierdo de la ecuación corresponde a la estandarización del lado derecho de la misma, y  $\pm$  representa la ambigüedad del signo, pudiendo definir el signo de los pesos externos de todo un bloque de variables para hacerlo coherente con la definición de la variable latente.



El *Paso 2* consiste en la estimación interna, en la cual cada variable latente es estimada considerando los vínculos con otras  $Q'$  variables latentes adyacentes:

$$\mathfrak{g}_q \propto \sum_{q'=1}^{Q'} e_{qq'} \mathfrak{v}_{q'} \quad (3.8)$$

donde  $\mathfrak{g}_q$  es la estimación interna estandarizada de la  $q$ -ésima variable latente  $\xi_q$ , y (en el esquema de centroide) cada peso interno ( $e_{qq'}$ ) presenta el mismo signo que la correlación entre la estimación externa  $\mathfrak{v}_q$  de la  $q$ -ésima variable latente y la estimación externa  $\mathfrak{v}_{q'}$  de la  $q'$  variable latente conectada con  $\mathfrak{v}_q$ .

Para la estimación de pesos internos, se puede utilizar uno de los tres tipos de esquemas descritos a continuación:

1. Esquema de Centroide: Toma el signo de la correlación entre la estimación externa  $\mathfrak{v}_q$  de la  $q$ -ésima variable latente y la estimación externa  $\mathfrak{v}_{q'}$ . Este esquema se adapta bien a casos en que las variables manifiestas están fuertemente correlacionadas (Wold, 1985).
2. Esquema de Factor (o factorial): Toma la correlación entre la estimación externa  $\mathfrak{v}_q$  de la  $q$ -ésima variable latente y la estimación externa  $\mathfrak{v}_{q'}$ . Este esquema se adecúa a casos donde la correlación entre variables manifiestas es débil (Lohmöller, 1989).

3. Esquema de Rutas (o estructural): Toma el coeficiente de regresión entre  $v_{q'}$  y  $v_q$  si  $v_q$  juega un rol de variable dependiente en la ecuación estructural específica, o toma el coeficiente de correlación en el caso que sea un predictor (Lohmöller, 1989). Este esquema, al considerar la dirección de las relaciones como son especificadas en el modelo predictivo de rutas, es el recomendado por Esposito Vinzi, Trinchera y Amato (2010).

El Paso 3 consiste en la actualización de los pesos externos  $w_{pq}$ . Para lo cual existen dos modos distintos de proceder:

1. Modo A: Cada peso externo  $w_{pq}$  es actualizado como el coeficiente de regresión en la regresión simple de la  $p$ -ésima variable manifiesta, del  $q$ -ésimo conjunto de variables ( $x_{pq}$ ) sobre la estimación interna de la  $q$ -ésima variable latente  $\xi_q$ . Como  $\xi_q$  es estandarizada, el peso externo genérico  $w_{pq}$  es obtenido como:

$$w_{pq} = \text{COV}(x_{pq}, \xi_q) \quad (3.9)$$

En el caso de que las variables manifiestas hayan sido estandarizadas la covarianza sería una correlación.

2. Modo B: El vector  $w_q$  de los pesos  $w_{pq}$  asociado a las variables manifiestas del  $q$ -ésimo grupo de variables es actualizado como el vector de coeficientes de regresión en la regresión múltiple de la

estimación interna de la  $q$ -ésima variable latente  $\mathfrak{g}_q$  sobre las variables manifiestas en  $X_q$  :

$$w_q = (X_q' X_q)^{-1} X_q' \mathfrak{g}_q \quad (3.10)$$

donde  $X_q$  comprende las  $P_q$  variables manifiestas  $x_{pq}$  previamente centrada y escalada a través de  $\sqrt{1/N}$  , para que la matriz de productos genere las matriz de varianzas-covarianzas.

La elección del modo de estimación de pesos externos se encuentra relacionada con la naturaleza del modelo de medida. Así para un Modelo de Medida Reflectivo el Modo A es más apropiado, mientras el Modo B es apropiado para un Modelo de Medida Formativo (Esposito Vinzi, Trinchera y Amato, 2010).

En el caso de un modelo constituido sólo por un bloque de variables, el Modo A lleva a los mismos resultados que la primera componente principal estandarizada en un Análisis de Componentes Principales. Sin un modelo interno, cualquier combinación lineal de las variables manifiestas es perfectamente explicada por las variables manifiestas en sí mismas.

El algoritmo del PLS-PM alterna los pasos de estimación interna y externa a través de iteraciones hasta lograr la convergencia.

- b. Etapa 2. Estimación de coeficientes de carga (modelo de medida):

En esta etapa se utilizan las puntuaciones de las variables latentes estimadas en la Etapa 1, para estimar los coeficientes de las relaciones internas y externas entre variables, a través regresiones de mínimos cuadrados ordinarios (OLS). Si se modela un grupo de variables reflectivo se utilizará un análisis de regresión simple (Ecuación (3.2)) para la estimación de cada carga ( $\lambda_{pq}$ ) y si el grupo de variables a modelar es formativo, se utilizará un análisis de regresión ordinaria múltiple (Ecuación (3.4)) para la estimación de  $\omega_{pq}$  (Martínez-Ruiz y Aluja-Banet, 2010).

La regresión OLS puede ser reemplazada por una regresión de mínimos cuadrados parciales (PLS) para la estimación de coeficientes de rutas cuando existan variables latentes fuertemente correlacionadas.

c. Etapa 3. Estimación de coeficientes de rutas (modelo estructural):

Esta etapa implica la estimación de los coeficientes de rutas ( $\beta_{qj}$ ) del modelo estructural a través de la Ecuación (3.1).

“Las puntuaciones de las variables latentes son obtenidas a través de la ecuación 3.6 hasta la convergencia. Así, PLS-PM provee una estimación directa de las puntuaciones individuales en las variables latentes como agregados de variables manifiestas que naturalmente involucran un error de medición. El precio de obtener estas puntuaciones es la inconsistencia de las estimaciones” (Esposito Vinzi, Trinchera y Amato, 2010, p. 55). Sin embargo, desde los primeros planteamientos sobre PLS-PM, se ha considerado que sus estimaciones son asintóticamente correctas, en el sentido de “*consistency at large*”, lo que implica que las estimaciones de PLS-PM tenderán a valores verdaderos en la medida que el tamaño muestral y el número de

indicadores tiendan al infinito (Gefen, Rigdon y Straub, 2011; Hui y Wold, 1982; Jöreskog y Wold, 1982; Lohmöller, 1989; McDonald, 1996, Wold, 1982a).

Una de las limitaciones más importantes de PLS-PM, es que “carece de un criterio global de optimización, por lo que no cuenta con una función de ajuste global para evaluar la bondad del modelo” (Esposito Vinzi, Trinchera y Amato, 2010, p. 56). Sin embargo, cuenta con de calidad local para validar cada parte del modelo: el modelo de medida (índice de comunalidad), el modelo estructural (índice de redundancia) y el modelo global (índice de bondad de ajuste). El índice de comunalidad mide la variabilidad de las variables manifiestas de cada bloque que es explicada por las puntuaciones de su variable latente, corresponde al promedio de correlaciones al cuadrado (cargas al cuadrado en el caso de variables manifiestas estandarizadas) entre cada variable manifiesta en el q-ésimo bloque y las puntuaciones de las variables latentes correspondientes. El índice de redundancia mide la porción de variabilidad de las variables manifiestas conectadas a la j-ésima variable latente explicada por las variables latentes directamente conectadas al bloque j; una medida de calidad global del modelo estructural es propuesta por Tenenhaus, Amato y Esposito Vinzi (2004), la cual es denominada índice de Bondad de Ajuste (GoF, “*Goodness of Fit*” index) y corresponde a la media geométrica del promedio de los índices de comunalidad y el valor promedio de coeficientes de determinación de las variables latentes endógenas del modelo. El índice de bondad de ajuste toma en cuenta el desempeño del modelo, tanto en el modelo de medida como en el modelo estructural, así entrega una única medida para el desempeño de predicción global del modelo (Esposito Vinzi, Trinchera y Amato, 2010).

Por último, se han desarrollado algunos temas avanzados en PLS-PM, tales como: modelos de componentes jerárquicos (Lohmöller, 1989), técnicas de segmentación, tal como Mínimos Cuadrados Parciales Mezcla Finita (FIMIX-PLS;

Ringle, Wende y Will, 2005b; Ringle Wende y Will, 2010), efectos no lineales (Rigdon, Ringle y Sarstedt, 2010) y guías para el análisis de efectos moderadores y de interacción (Henseler y Chin, 2010). Estos avances han expandido la utilidad de PLS-PM como herramienta en marketing y ciencias sociales (Hair, y cols. 2012).

## Capítulo 4. Análisis Generalizado de Componentes Estructurales (GSCA)

En la Figura 5 se incorpora la notación propia de GSCA utilizada por Hwang y Takane (2004).

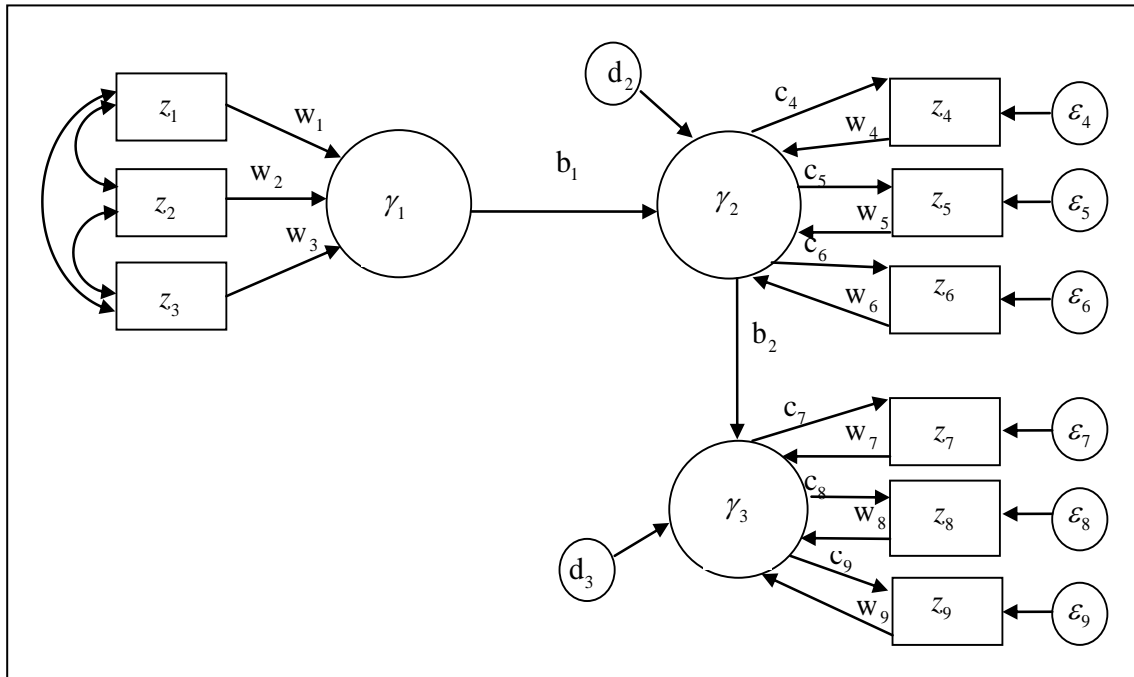


Figura 5. Ejemplo de Diagrama de rutas de un Análisis Generalizado de Componentes Estructurales.

GSCA, al igual que el PLS-PM, constituye una aproximación basada en componentes, es decir que cada bloque, componente o variable latente  $\gamma_i$  es calculada como una suma ponderada de las variables manifiestas con las que se encuentra vinculada (Hwang y Takane, 2004; Hwang y cols., 2010a):

$$\gamma_i = \mathbf{Wz}_i \quad (4.1)$$

donde  $\mathbf{W}$  es una matriz de pesos asignados a las variables manifiestas asociadas a cada componente; y  $\mathbf{Z}_i$  denota un vector de variables observadas de un respondiente  $i$  ( $i=1, \dots, N$ ).

Además, en la especificación de un modelo, GSCA incluye dos ecuaciones más. Una para el Modelo Externo (Modelo de Medida) y otra para el Modelo Interno (Modelo Estructural). El modelo de medida incluye las relaciones entre las variables manifiestas y las variables latentes, según:

$$\mathbf{z}_i = \mathbf{C}\boldsymbol{\gamma}_i + \boldsymbol{\varepsilon}_i \quad (4.2)$$

donde  $\mathbf{C}$  es la matriz de saturaciones que relacionan las variables latentes con las variables manifiestas; y  $\boldsymbol{\varepsilon}_i$  es el vector con los residuales para  $\mathbf{Z}_i$ . El modelo estructural incluye las relaciones entre las variables latentes, según:

$$\boldsymbol{\gamma}_i = \mathbf{B}\boldsymbol{\gamma}_i + \mathbf{d}_i \quad (4.3)$$

donde  $\mathbf{B}$  es una matriz de coeficientes de ruta que conectan las variables latentes entre sí; y  $\mathbf{d}_i$  es un vector de residuales para  $\boldsymbol{\gamma}_i$ .

El GSCA se deriva de la combinación de las tres ecuaciones en una sola, de acuerdo con:

$$\mathbf{ZV} = \mathbf{ZWA} + \mathbf{E} \quad (4.4)$$



donde  $\mathbf{Z}$  denota una matriz  $N \times J$  de variables observadas;  $\mathbf{V}$  es una matriz  $J \times T$  de pesos asociados a las variables endógenas;  $\mathbf{W}$  es una matriz  $J \times D$  de pesos asociados a las variables exógenas;  $\mathbf{A}$  es una súper matriz  $D \times T$ , esto es  $\mathbf{A} = [\mathbf{C}, \mathbf{B}]$ ; y  $\mathbf{E}$  es una matriz de residuales.

Los parámetros desconocidos  $\mathbf{V}$ ,  $\mathbf{W}$  y  $\mathbf{A}$  son estimados de forma que la suma de cuadrados de los residuales  $\mathbf{E}$  sea lo más pequeña posible. Lo cual es equivalente a minimizar el siguiente criterio de mínimos cuadrados:

$$\phi = \sum_{i=1}^N (\mathbf{Vz}_i - \mathbf{AWz}_i)' (\mathbf{Vz}_i - \mathbf{AWz}_i) \quad (4.5)$$

Lo anterior constituye uno de los principales aportes de GSCA, frente a PLS-PM, ya que le permite mantener las propiedades de PLS-PM en cuanto a la ausencia de requerimientos distribucionales para los datos y la ausencia de problemas de convergencia, permitiendo además la estimación de un índice global de ajuste (FIT), el cual es la proporción de varianza total de todas las variables endógenas explicada por la especificación particular del modelo. Por otra parte, GSCA permite ajustar un mismo modelo a más de un grupo, pudiendo imponer restricciones entre los grupos (Hwang y Takane, 2004; Hwang y cols., 2010a).

El algoritmo para minimizar el criterio  $\phi$  alterna dos pasos hasta la convergencia. En el primer paso  $\mathbf{A}$  es actualizada, manteniendo fijas  $\mathbf{V}$  y  $\mathbf{W}$ . En el segundo paso  $\mathbf{V}$  y  $\mathbf{W}$  son actualizadas, manteniendo fija  $\mathbf{A}$ . Este algoritmo es denominado mínimos cuadrados alternantes (ALS; De Leeuw, Young y Takane, 1976) y puede ser considerado como un tipo especial del algoritmo de punto fijo de PLS-PM.

GSCA mide el ajuste global de un modelo hipotético, a través de la varianza total de todas las variables endógenas explicada por las predicciones del modelo especificado (FIT):

Si,

$$SS(\mathbf{X}) = tr(\mathbf{X}\mathbf{X}) \quad (4.6)$$

Entonces,

$$FIT = 1 - \frac{SS(\mathbf{ZV} - \mathbf{ZWA})}{SS(\mathbf{ZV})} \quad (4.7)$$

“Este índice de ajuste se encuentra en un rango de 0 a 1, donde mayores valores del índice de ajuste implican que mayor varianza de las variables endógenas es explicada por el modelo” (Hwang y Takane, 2004, pp. 88). “La medida FIT es una función de la suma de residuales cuadrados, que resume las discrepancias entre el modelo y los datos. Sin embargo, el índice FIT se encuentra influenciado por la complejidad del modelo, así a mayor cantidad de parámetros, mayor es el valor de FIT” (Hwang y cols., 2010a, pp. 701); por lo cual Hwang, DeSarbo y Takane (2007) proponen una corrección por grados de libertad, generando el índice FIT corregido (AFIT):

$$AFIT = 1 - (1 - Fit) \left( \frac{d_0}{d_1} \right) \quad (4.8)$$

donde  $d_0$  son los grados de libertad del modelo nulo,  $d_1$  son los grados de libertad del modelo puesto a prueba. El modelo que maximiza AFIT es considerado el más apropiado cuando se comparan modelos.

Dado el reciente desarrollo de la técnica de MEE GSCA, no se han desarrollado muchas extensiones para su aplicación. Sin embargo, Hwang, Ringo Ho y Lee (2010) plantean un procedimiento para analizar distintos tipos de interacciones entre variables latentes con GSCA.

Recientemente, Jörg Henseler (2012), planteó críticas importantes al estudio realizado por Hwang y cols. (2010a). La primera de ellas se refiere a la inclusión de dos factores al mismo tiempo, por lo que los efectos observados sólo podrían ser atribuidos a ambos factores en conjunto. Los autores “combinaron el modelo de medida (correctamente especificado versus incorrectamente especificado) e incluyeron un parámetro estructural adicional (fijado a 0 versus libremente estimado)... por lo que no queda claro si GSCA es inmune a la incorrecta especificación del modelo de medida o es capaz de reconocer un efecto de 0 como tal” (Henseler, 2012, p. 407). La segunda crítica es la presencia de un error en el algoritmo de GSCA, el que utiliza variables observadas estandarizadas y variables latentes normalizadas, por lo que maximiza un criterio distinto de FIT. Henseler (2012) plantea que este error tiene consecuencias para todos los estudios empíricos que han aplicado GSCA, haciendo inválidos sus hallazgos empíricos. Hwang y cols. (2010b) reconocen este error, publicando una errata y entregando una breve descripción de los resultados obtenidos con la corrección del error, normalizando tanto las variables observadas como las variables latentes y señalando que el error no cambia las conclusiones ni recomendaciones de su estudio. Además en noviembre de 2010 se realizó la corrección en el programa informático GeSCA (Hwang, 2009), desarrollado para

aplicar la técnica de MEE GSCA. La que es informada en la página web del programa informático ([www.sem-gesca.org](http://www.sem-gesca.org)). La tercera crítica se refiere a la generalización de los resultados a otros modelos. Los autores “no controlaron interferencias potenciales de elementos constitutivos del modelo poblacional, tales como el signo de los coeficientes o la existencia de cargas cruzadas” (Henseler, 2012, p. 407). Por último, Henseler (2012) plantea que GSCA tiende a proveer de estimaciones inconsistentes, sugiriendo equipar a GSCA con alguna forma de corrección para su atenuación.

## Capítulo 5. Comparación de ACOV, PLS-PM y GSCA

En los capítulos 2, 3 y 4, se evidencia que las técnicas de MEE comparten ciertas características en cuanto a la representación gráfica, por ejemplo utilizan círculos para representar las variables latentes, cuadrados o rectángulos para representar las variables manifiestas, líneas curvas para representar covarianza entre variables y líneas rectas para representar relaciones causales entre variables. Sin embargo, difieren en la mayoría de las letras utilizadas para designar variables y parámetros de los modelos. En la Tabla 1 se presentan las letras utilizadas por ACOV, según Bollen (1989); PLS-PM, según Espósito Vinzi, Trinchera y Amato (2010); y GSCA, según Hwang y Takane (2004) para designar variables y parámetros.

Tabla 1. Notación para variables y parámetros de las tres técnicas de MEE.

	ACOV	PLS-PM	GSCA
Variable exógena	$\xi$	$\xi$	$\gamma$
Variable endógena	$\eta$	$\xi$	$\gamma$
Variable manifiesta asociada a variable exógena	$x$	$x$	$z$
Variable manifiesta asociada e variable endógena	$y$	$x$	$z$
Error de medición de variable manifiesta asociada a variable exógena	$\delta$	$\varepsilon$	$\varepsilon$
Error de medición de variable manifiesta asociada a variable endógena	$\varepsilon$	$\varepsilon$	$\varepsilon$
Error de predicción de variable latente endógena	$\zeta$	$\zeta$	$d$
Error de predicción de variable latente que recibe efecto de variables manifiestas formativas	$\zeta$	$\delta$	-
Efecto de variable exógena sobre variable endógena	$\gamma$	$\beta$	$b$
Efecto de variable endógena sobre variable endógena	$\beta$	$\beta$	$b$
Efecto de variable latente sobre variable manifiesta	$\lambda$	$\lambda$	$c$
Efecto de variable manifiesta sobre variable latente	$\gamma$	$w$	$w$

Cabe señalar que no existe acuerdo sobre la nomenclatura de las técnicas. Así otros autores utilizan una notación distinta a la presentada en la Tabla 1. Por ejemplo, Mulaik (2009) propone variaciones a la notación utilizada para ACOV, denominando con la letra griega  $\varepsilon$  a todos los errores, tanto los errores de medición (de variables manifiestas asociadas a variables latentes endógenas y exógenas) y de predicción de las variables latentes; además utiliza la letra griega  $\delta$  para denominar los parámetros que representan el peso de los errores de medición y predicción, entre otras propuestas en la notación. Por su parte, originalmente Wold (1980) utilizó la letra griega  $\pi$  para denominar las cargas (“loadings”) de las variables latentes sobre las variables manifiestas en PLS-PM. En cuanto a GSCA, no se han encontrado variaciones a la nomenclatura propuesta de Hwang y Takane (2004). Aún cuando en el estudio de Hwang y cols. (2010a) utilizan  $\eta$  para denominar a las variables latentes en el diagrama de rutas. Sin embargo, se atribuye dicha variación en la notación en el diagrama de rutas, a un intento de ser fieles a la nomenclatura utilizada por los autores que utilizaron el modelo originalmente (Paxton, Curran, Bollen, Kirby y Chen, 2001). Ya que en la descripción del procedimiento del GSCA utilizan la nomenclatura original de Hwang y Takane (2004), utilizando  $\gamma$  para denominar las variables latentes.

Respecto a los métodos para la estimación de parámetros, ACOV puede utilizar distintos métodos (p.e. ML, ULS, GLS y ADF) aunque suele utilizarse ML, por ser un método de estimación considerado robusto y por encontrarse como método de estimación por defecto en los programas estadísticos. Por su parte PLS-PM estima los parámetros de un modelo a través del algoritmo de punto fijo (FP; Lyttkens, 1968, 1973; Wold, 1965, 1981); y GSCA utiliza la variación del algoritmo de punto fijo, denominada algoritmo de mínimos cuadrados alternante (ALS; De Leeuw, Young y Takane, 1976).

Las primeras comparaciones entre técnicas de MEE (Areskoug, 1982; Chin, 1995; Fornell y Bookstein, 1982; Jöreskog y Wold, 1982; Reinartz, Haenlein y Henseler, 2009; Wold, 1982a) consideran a ACOV y PLS-PM, por ser las primeras en ser desarrolladas. Las comparaciones concuerdan en que ambas técnicas son complementarias más que competitivas, destacando que el propósito de ACOV es estudiar la estructura de la relación entre variables, reflejada en la matriz de varianzas y covarianzas, y PLS-PM tiene como objetivo principal maximizar la varianza explicada de las variables dependientes del modelo. Así la utilización de PLS-PM es más apropiada para situaciones en las cuales el investigador busca predecir las variables latentes dependientes del modelo o identificar de manera exploratoria las relaciones entre ellas; mientras que la utilización de ACOV debe ser escogida cuando el propósito recae en confirmar relaciones asumidas teóricamente (Haenlein y Kaplan, 2004; Reinartz, Haenlein y Henseler, 2009). Entre las diferencias se encuentra que ACOV suele asumir la distribución normal conjunta (cuando se utiliza ML como método de estimación de parámetros), mientras que PLS-PM no realiza supuestos de distribución de los datos. Las estimaciones de ACOV en muestras grandes con distribución normal multivariante, han resultado ser eficientes y más precisas que las estimaciones de PLS-PM (Fornell y Bookstein, 1982). Las estimaciones del PLS-PM no presentan problemas en la identificación del modelo, son eficientes en muestras pequeñas, aun estimando modelos complejos, y logran la convergencia más rápidamente (Fornell y Bookstein, 1982; Jöreskog y Wold, 1982). Por otra parte, se plantea que las estimaciones de parámetros de PLS-PM presentan una mayor eficiencia pero también un mayor sesgo, frente a las estimaciones del ACOV; tendiendo a subestimar los parámetros del modelo estructural y a sobreestimar los parámetros del modelo de medida (Chin, 1995; Hulland, Ryan y Rayner, 2010; Vilares, Almeida y Coehlo, 2010).

La diferencia clave de PLS-PM respecto de ACOV es la utilización de una definición explícita de las variables latentes, como un combinación ponderada de sus

indicadores por parte del primero (Wold, 1982b), de manera similar a como se define el modelo de Componentes Principales (Chin, 1995). Este hecho ha favorecido que se le denomine MEE basado en componentes (Haenlein y Kaplan, 2004; Tenenhaus, 2008), y a veces también se le ha llamado MEE basado en la varianza (Lohmöller, 1989; Wold, 1982a). Por su parte, ACOV (muy enraizada en la filosofía de los modelos de análisis factorial) propone la definición de los indicadores como una combinación de las variables latentes del modelo y, además, plantea la necesidad de estimar las communalidades como sucede en el modelo de Factor Común (Chin, 1995), lo que ha conllevado que se le denomine MEE basado en factores, a veces también llamado MEE basado en covarianzas (Jöreskog, 1978; Tenenhaus, 2008).

En 1998, Chin (1998b) planteó que una asunción para el análisis de MEE es que los ítems o indicadores usados para medir una variable latente son reflectivos en naturaleza, es decir, que reflejan la magnitud del constructo latente que pretenden medir. Esto plantea que para técnicas de MEE tales como ACOV, que tienen en cuenta las covarianzas entre las medidas, la inclusión de medidas formativas será problemática. Sin embargo, MacCallum y Browne, 1993 proponen una estrategia para evitar problemas de identificación al estimar modelos con variables causales con ACOV. La que consiste en fijar a 0 la varianza del error de predicción de la variable latente afectada por las variables formativas. Además, Williams, Edwards y Vandenberg (2003) proponen una estrategia para la especificación de un modelo con variables manifiestas formativas con ACOV, la que consiste en reespecificar las variables manifiestas formativas como variables latentes exógenas con una sola variable manifiesta, fijando sus cargas a 1 y fijando a 0 la varianza de sus errores de medida.

Ringle, Götz, Wetzels y Wilson (2009), estudian distintos métodos de estimación de ACOV y distintos esquemas de ponderación de PLS-PM con un modelo



que incluye variables manifiestas formativas continuas, en dos condiciones de asimetría (0 y 2). Los resultados apuntan que frente a un modelo con variables formativas las estimaciones de ACOV con Máxima Verosimilitud (ML) son más precisas y robustas que las estimaciones con los métodos de Mínimos Cuadrados No Ponderados (ULS), Mínimos Cuadrados Generalizados (GLS) y Asintóticamente Libre de Distribución (ADF). Por otra parte comprueban que los tres esquemas de ponderación de PLS-PM: centroide, factor y ruta, los que serán descritos en el capítulo 3, entregan las mismas estimaciones para los parámetros, lo cual concuerda con lo planteado por Tenenhaus, Esposito Vinzi, Chatelin, y Lauro, (2005). En cuanto a la comparación de ambas técnicas de MEE evidencian que las estimaciones de los parámetros formativos del modelo de medida tienden a ser sobreestimados por ACOV y subestimados por PLS-PM, mientras que los parámetros reflectivos del modelo de medida tienden a ser sobreestimados por PLS-PM y subestimados por ACOV. Finalmente los parámetros del modelo estructural tienden a ser sobrestimados por ACOV y subestimados por PLS-PM. Las limitaciones del estudio incluyen la utilización de un solo tamaño muestral ( $n=300$ ) y no se incorporan variaciones en la especificación del modelo.

En los estudios que comparan ACOV con PLS-PM se suelen utilizar datos simulados continuos (p.e. Fornell y Bookstein, 1982; Reinartz, Haenlein y Henseler, 2009). Aunque actualmente se han realizado estudios con datos en escala ordinal (p.e. Barroso, Cepeda y Roldán, 2010; Hulland, Ryan y Rayner, 2010; y Vilares, Almeida y Coehlo, 2010). Estos estudios utilizan ACOV con el método de estimación de Máxima Verosimilitud (ML), por ser el método de estimación más utilizado, ser considerado robusto y ser método de estimación programado por defecto en la mayoría de programas informáticos, a pesar del planteamiento de Elosua y Zumbo (2008), Ferrando y Lorenzo Seva (1993), Flora y Curran (2004) y Fornell y Bookstein (1982), sobre la conveniencia de utilizar el método de estimación de Mínimos Cuadrados No

Ponderados (ULS; Olsson, 1979) sobre la matriz de correlaciones policóricas en el caso de utilizar variables manifiestas en escala ordinal; y los hallazgos de Forero, Maydeu-Olivares y Gallardo-Pujol (2009), quienes comparan la precisión en las estimación de parámetros de los métodos de estimación Mínimos Cuadrados Ponderados Diagonales (DWLS) y ULS, en datos con escala ordinal, encontrando una mayor precisión al utilizar el método de estimación de ULS sobre la matriz de correlaciones policóricas.

Tenenhaus, Mauger y Guinot (2010) utilizan el método de estimación ULS para ACOV. En su estudio proponen una metodología para estimar el efecto de grupo en el modelo estructural, cuando este no afecta el modelo de medida. Sin embargo, la aplicación la realizan sobre un modelo con variables manifiestas dicotómicas. Los autores utilizan ACOV y PLS-PM en cuatro pasos, para la estimación por lo que no comparan ambas técnicas, sino que las integran en una metodología. Primero utilizan ACOV, con el programa AMOS para calcular los pesos de las variables manifiestas, los cuales han sido restringidos como iguales en los distintos grupos. Luego a través de PLS-PM calculan las variables latentes utilizando los pesos determinados por AMOS. Posteriormente estudian la interacción a través de un análisis de covarianza relacionando el bloque de la variable latente de respuesta al bloque de la variable latente predictora, el efecto de grupo y la interacción de cruzar el bloque de la variable latente predictora y el efecto de grupo. Finalmente, utilizan el remuestreo (“bootstrapping”) de AMOS para producir los intervalos de confianza de los coeficientes de regresión de las ecuaciones estructurales.

En el 2004, Hwang y Takane presentan formalmente la técnica de MEE GSCA, la cual, como se explicita en su nombre, constituye una aproximación basada en componentes. Un trabajo destacable, es el realizado por Hwang y cols. (2010a), quienes en su estudio de simulación comparan la eficiencia de ACOV, PLS-PM y

GSCA con datos simulados continuos, en un modelo con variables manifiestas reflectivas que incluye efectos cruzados. Los autores describen GSCA, destacando que aunque comparte con PLS-PM la definición explícita de las variables latentes, como una combinación ponderada de sus indicadores, se diferencia de PLS-PM al contar con una única ecuación que integra el submodelo de medida y el submodelo estructural (tal como ACOV), lo que junto con la utilización de una función de optimización global le permite proveer de una medida de ajuste global del modelo, manteniendo las ventajas de PLS-PM como técnica de MEE basada en componentes.

Hwang y cols. (2010a) encontraron que la única condición determinante en la recuperación de parámetros fue la especificación correcta del modelo. En concreto, cuando el modelo se encontraba correctamente especificado, ACOV recuperó los parámetros sin sesgo y recuperó mejor los parámetros que PLS-PM y GSCA. Sin embargo, cuando el modelo se encontraba incorrectamente especificado GSCA recuperó mejor los parámetros. En cuanto al sesgo, ACOV sobreestimó los parámetros del modelo estructural en el límite del rango aceptable (10%) y del modelo de medida fuera del rango aceptable; y PLS-PM subestimó los parámetros del modelo estructural dentro del rango aceptable y sobreestimó los parámetros del modelo medida fuera del rango aceptable.

En la Tabla 2 se presenta la comparación de las tres técnicas de MEE realizada por Hwang y cols. (2010a). Al respecto señalan que, en cuanto a la especificación del modelo, tanto GSCA como PLS-PM definen las variables latentes como componentes, a diferencia de ACOV, en el cual las variables latentes son equivalentes a factores comunes. Lo que implica que las variables latentes en el ACOV son aleatorias; y en GSCA y PLS-PM son fijas. Esto determina la especificación de diferentes conjuntos de parámetros del modelo para las variables latentes (p.e. media de factores y varianzas error en ACOV y pesos de los componentes en GSCA y

PLS-PM). En cuanto a los datos de entrada requeridos para la aplicación de las técnicas de MEE, ACOV utiliza la matriz de varianzas-covarianzas o la matriz de correlaciones entre las variables observadas; y GSCA y PLS-PM utilizan los datos en bruto a nivel individual. Por su parte GSCA, al igual que ACOV cuenta con una única ecuación que integra el submodelo de medida y el submodelo estructural, lo que le permite proveer de una medida de ajuste global del modelo. En cuanto al método de estimación, ACOV suele utilizar máxima verosimilitud, el cual requiere el cumplimiento del supuesto de normalidad multivariante; y PLS-PM y GSCA utilizan mínimos cuadrados que no poseen el supuesto de normalidad multivariante. Los autores proponen GSCA como una alternativa a PLS-PM y GSCA, en condiciones de un modelo mal especificado, instancia que consideran puede ser habitual en investigaciones aplicadas.

Tabla 2. Comparación de las tres técnicas de MEE.

	<b>ACOV</b>	<b>PLS-PM</b>	<b>GSCA</b>
<b>Propósito del Análisis</b>	<i>Estimar parámetros estructurales</i>	<i>Predecir variables dependientes</i>	<i>Predecir variables dependientes</i>
<b>Especificación del modelo</b>			
Variables latentes	Factores	Componentes	Componentes
Cantidad de ecuaciones	Una	Dos	Una
Parámetros del modelo	Cargas, coeficientes de ruta, varianzas error, medias y varianzas de factores	Cargas, coeficientes de ruta, pesos de componentes	Cargas, coeficientes de ruta, pesos de componentes
<b>Estimación de parámetros</b>			
Datos de entrada	Covarianzas / correlaciones	Datos brutos individuales	Datos brutos individuales
Método de estimación	Máxima verosimilitud (principalmente)	Mínimos cuadrados (algoritmo de punto fijo, FP)	Mínimos cuadrados (algoritmo alternante, ALS)
Función de optimización global	Sí	No	Sí
Supuesto de normalidad	Requerido para máxima verosimilitud	No requerido	No requerido
Medidas de ajuste del modelo	Global y local	Local	Global y local

Para profundizar en la comparación de las técnicas de MEE: ACOV, PLS-PM y GSCA. A continuación, en la parte empírica de este trabajo, se presentan dos estudios, el primero de ellos busca evaluar la eficiencia y sesgo de ACOV y PLS-PM en la recuperación de parámetros con datos simulados con escala ordinal, en un modelo con variables manifiestas reflectivas, considerando distintas condiciones de: a) asimetría, b) especificación del modelo y c) tamaño muestral. El segundo busca comparar ACOV, PLS-PM y GSCA respecto de su eficiencia y sesgo en la recuperación de parámetros en un modelo que incluye variables manifiestas reflectivas y formativas, con datos simulados con escala ordinal, considerando distintos niveles de asimetría, especificación del modelo y tamaño muestral; superando algunas de las limitaciones presentes en el trabajo de Hwang y cols. (2010a) y señaladas por los propios autores. Este estudio limita la cantidad de condiciones y de réplicas por condición para incorporar la aplicación de GSCA, dada la limitación para la automatización de la aplicación de esta técnica de MEE presentada por el único programa informático disponible y reconocida por uno de los autores del mismo (Hwang, 2011). Se destaca la utilización de variables manifiestas con escala ordinal, dada su amplia utilización en distintas áreas de la psicología; la incorporación de variables manifiestas formativas considerando su utilización en el ámbito de las ciencias económicas; y la utilización del método de estimación ULS para ACOV, el que ha demostrado ser el más eficiente cuando se utilizan datos con escala ordinal y cuando se utiliza un modelo que incluyen variables manifiestas formativas.



## **PARTE EMPÍRICA**





# **Estudio 1: Eficiencia y sesgo en la recuperación de parámetros, con datos ordinales simulados en un modelo con variables manifiestas reflectivas, del Análisis de Estructura de Covarianza y el Modelo de Rutas mediante Mínimos Cuadrados Parciales**

## **Introducción**

Al ser las primeras técnicas de MEE desarrolladas, ACOV (Jöreskog, 1970a, 1970b) y PLS-PM (Wold, 1977, 1980) han sido compradas en diversas ocasiones (Areskoug, 1982; Chin, 1995; Fornell y Bookstein, 1982; Jöreskog y Wold, 1982; Reinartz, Haenlein y Henseler, 2009; Wold, 1982a). Las comparaciones apuntan a que ambas técnicas son más complementarias que competitivas, destacando que el propósito de ACOV es estudiar la estructura de las relaciones entre variables, reflejada en la matriz de varianzas y covarianzas; mientras que PLS-PM tiene como objetivo principal maximizar la varianza explicada de las variables dependientes del modelo. Así, la utilización de PLS-PM es más apropiada para situaciones en las cuales el interés fundamental del investigador es predecir las variables latentes del modelo o identificar de manera exploratoria las relaciones existentes entre ellas; mientras que la utilización de ACOV debe ser escogida cuando el propósito recae en confirmar relaciones teóricamente asumidas (Haenlein y Kaplan, 2004; Reinartz, Haenlein y Henseler, 2009). Ambas técnicas de MEE comparten las convenciones de representación gráfica del modelo y la formulación lineal de las relaciones entre las variables manifiestas y las variables latentes. Entre las diferencias se encuentra que ACOV asume la distribución normal conjunta, mientras que PLS-PM no realiza supuestos distribucionales. Las estimaciones de ACOV en muestras grandes con distribución normal multivariante, han resultado ser más eficientes que las estimaciones de PLS-PM (Fornell y Bookstein, 1982), mientras que las estimaciones

del PLS-PM no presentan problemas en la identificación del modelo, son eficientes en muestras pequeñas, aun estimando modelos complejos, y logran la convergencia más rápidamente (Fornell y Bookstein, 1982; Jöreskog y Wold, 1982). Por otra parte, se plantea que las estimaciones de parámetros de PLS-PM presentan una mayor eficiencia y sesgo, frente a las estimaciones del ACOV; tendiendo a subestimar los parámetros del modelo estructural y a sobreestimar los parámetros del modelo de medida (Chin, 1995; Hulland, Ryan y Rayner, 2010).

La diferencia clave de PLS-PM respecto de ACOV es la utilización de una definición explícita de las variables latentes, como un combinación ponderada de sus variables manifiestas por parte del primero (Wold, 1982b), de manera similar a como se define el modelo de Componentes Principales (Chin, 1995), por lo que es parte de la aproximación de MEE basada en componentes, a veces también llamada aproximación basada en varianzas. Por su parte, ACOV propone la definición de las variables manifiestas como una combinación de las variables latentes y, además, plantea la necesidad de estimar las comunalidades como en el modelo de Factor Común (Chin, 1995), por lo que es parte de la aproximación de MEE basado en covarianzas, a veces también llamado MEE basado en factores.

Un trabajo destacable, es el realizado por Hwang, Malhotra, Kim, Tomiuk y Hong (2010a). En su estudio, realizaron simulaciones para comparar la precisión en la recuperación de parámetros de ACOV, PLS-PM y GSCA con variables manifiestas continuas, en un modelo que incluye efectos cruzados. Los autores utilizaron distintos niveles de asimetría (0 y 1,25), distintos tamaños muestrales (100, 200, 300, 400 y 500) y dos niveles de especificación del modelo (bien especificado y mal especificado). Encontrando que la única condición determinante en la recuperación de parámetros fue la especificación del modelo. En concreto, considerando las técnicas PLS-PM y ACOV, cuando el modelo se encontraba correctamente especificado, ACOV recuperó

los parámetros sin sesgo y recuperó mejor los parámetros que PLS-PM. Sin embargo, cuando el modelo se encontraba incorrectamente especificado PLS-PM recuperó mejor los parámetros, aún cuando ambas técnicas presentaron sesgo en sus estimaciones. ACOV sobreestimó los parámetros del modelo estructural en el límite del rango aceptable (10%) y del modelo de medida fuera del rango aceptable; y PLS-PM subestimó los parámetros del modelo estructural dentro del rango aceptable y sobreestimó los parámetros del modelo medida fuera del rango aceptable.

Los estudios que comparan ACOV con PLS-PM, suelen utilizar datos simulados continuos (p.e. Fornell y Bookstein, 1982; Reinartz, Haenlein y Henseler, 2009). Aunque actualmente se han realizado estudios con datos en escala ordinal (p.e. Barroso, Cepeda y Roldán, 2010; Hulland, Ryan y Rayner, 2010), estos estudios utilizan ACOV con el método de estimación de Máxima Verosimilitud (ML), por ser el método de estimación más utilizado, por ser considerado robusto y por ser el método de estimación programado por defecto en la mayoría de programas informáticos. Sin embargo, Elosua y Zumbo (2008), Ferrando y Lorenzo Seva (1993), Flora y Curran (2004) y Fornell y Bookstein (1982), sugieren utilizar el método de estimación de Mínimos Cuadrados No Ponderados (ULS) sobre la matriz de correlaciones policóricas, cuando se utilizan variables manifiestas con escala ordinal. Por su parte, Forero, Maydeu-Olivares y Gallardo-Pujol (2009), comparan los métodos de estimación Mínimos Cuadrados Ponderados Diagonales (DWLS) y ULS, encontrando una mayor precisión en las estimaciones de parámetros, al utilizar el método de estimación de Mínimos Cuadrados No Ponderados (ULS) sobre la matriz de correlaciones policóricas.

En el 2010, Tenenhaus, Mauger y Guinot, proponen una metodología de cuatro pasos para estimar el efecto de grupo en el modelo estructural, cuando este no afecta al modelo de medida. Los autores incluyen las técnicas de MEE: ACOV y PLS-PM

utilizando el método de estimación ULS para ACOV. Sin embargo, las variables manifiestas del modelo tienen escala dicotómica.

El presente estudio persigue profundizar en la comparación de las técnicas de MEE: ACOV y PLS-PM, utilizando condiciones extremas de asimetría y tamaño muestral, utilizando variables manifiestas con escala ordinal, condiciones que se consideran habituales en el ámbito de investigación aplicada en las ciencias del comportamiento y de la salud. Aplicando ACOV con el método de estimación ULS al ser el método de estimación más eficiente para las estimaciones con variables con escala ordinal.

## **Método**

Se diseñó un estudio de recuperación de parámetros mediante simulación de muestras sintéticas. Se consideró pertinente valorar el efecto conjunto de 4 factores: Dos factores intrasujeto: (1) Técnica de MEE: ACOV y PLS-PM; y (2) Especificación del modelo: Más especificado (se especifica el modelo poblacional) y menos especificado (se libera el parámetro  $\gamma_{21}$ , para ser estimado, el que se ha fijado a 0 en la generación de las réplicas), y dos factores intersujeto: (1) Nivel de asimetría: 0 (nula), 1.25 (media) y 2 (alta); y (2) Tamaño de la muestra: 50 (pequeño), 100 (medio), 300 (grande) y 500 (muy grande). Se planteó un diseño completo (2x2x3x4) para la combinación de los niveles de los factores, generando muestras independientes para cada combinación de niveles.

El modelo poblacional utilizado (Figura 6) incluye tres variables latentes (una exógena y dos endógenas) y nueve variables manifiestas reflectivas (tres por cada variable latente). La variable latente exógena ( $\xi$ ) tiene un efecto directo sobre una de

las variables latentes endógenas ( $\eta_1$ ), la cual a su vez tiene un efecto directo sobre la otra variable latente endógena ( $\eta_2$ ). Este modelo es una variación del modelo utilizado por Bollen, Kirby, Curran, Paxton y Chen (2007) y Hwang y cols. (2010a) para evaluar la recuperación de parámetros con distintos métodos de estimación utilizando variables manifiestas continuas. Se excluyeron los efectos cruzados (cross-loadings) de las variables latentes hacia las variables manifiestas, para considerar una estructura lo más simple posible. Además se utilizan variables tipificadas.

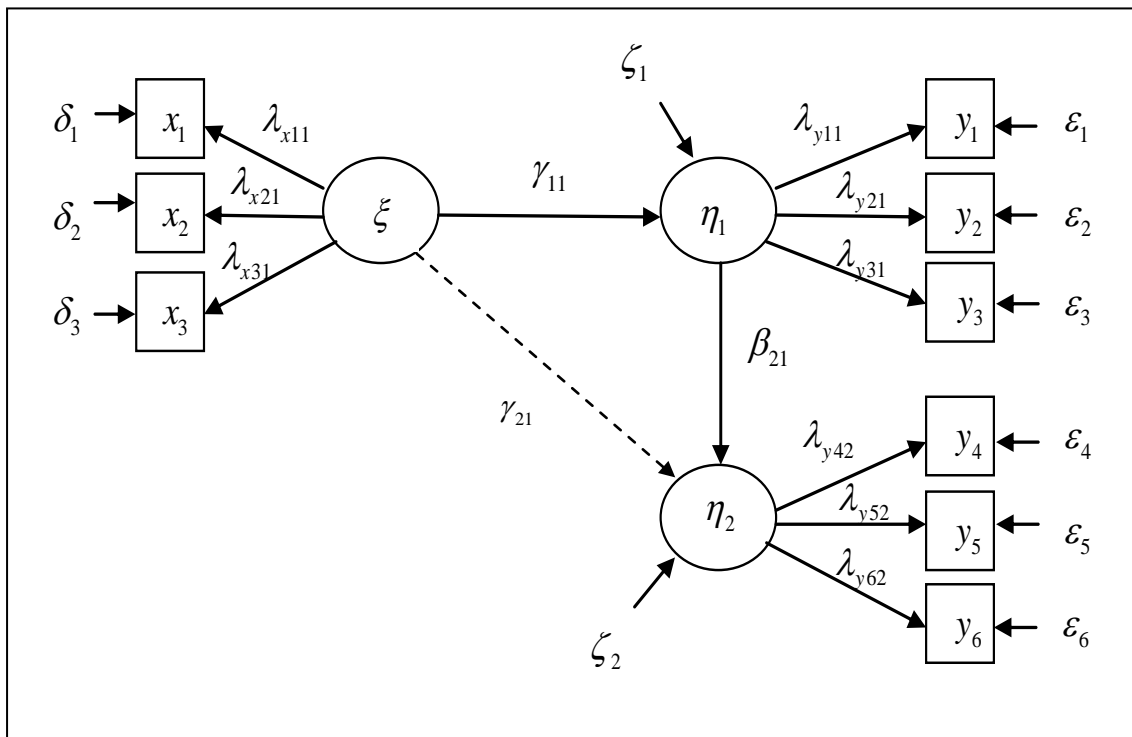


Figura 6. Diagrama de rutas del modelo poblacional, con variables manifiestas reflectivas.

\* La línea segmentada representa el parámetro que se agrega en el modelo menos especificado.

### Simulación de los datos

Se realizó una simulación Monte Carlo proactiva (Marcoulides, 2006). La que consiste en generar datos a partir de una población con valores de parámetros

hipotéticos, extrayendo muestras aleatorias repetidamente para proveer de estimaciones de los parámetros.

Se utilizó el programa Matlab 7.9.0 (Anexo A), para generar 500 muestras para cada combinación de niveles de las condiciones experimentales inter-grupos, simulando un total de 6000 muestras (3 niveles de asimetría  $\times$  4 tamaños de la muestra  $\times$  500 réplicas), siguiendo el procedimiento que se describe a continuación:

1.- Se definen los valores de las cargas factoriales, con  $\lambda = 0,70$  para todas las variables manifiestas y se construyen las matrices  $\Lambda_x$  y  $\Lambda_y$  según el modelo teórico especificado.

2.- Se definen los valores de las matrices  $\Phi = [1]$ ,  $\Gamma = \begin{bmatrix} 0,60 \\ 0 \end{bmatrix}$ ,  $\mathbf{B} = \begin{bmatrix} 0 & 0 \\ 0,60 & 0 \end{bmatrix}$  y  $\Psi = \begin{bmatrix} 0,64 & 0 \\ 0 & 0,64 \end{bmatrix}$ ; los valores de las varianzas de los errores de predicción de las

variables latentes endógenas han sido calculados para que el valor de las varianzas de las variables latentes endógenas sea igual a 1.

3.- Se construyen las matrices, que conformarán la matriz  $\Sigma(\theta)$  (Ecuación 1) reproducida en el modelo (Bollen, 1989). Las matrices  $\Sigma_{xx}(\theta)$  y  $\Sigma_{yy}(\theta)$  incluyen las matrices diagonales  $\Theta_\delta$  y  $\Theta_\varepsilon$  respectivamente. Las últimas complementan los valores de las diagonales de las primeras para que éstas contengan unos.

$$\Sigma(\theta) = \begin{bmatrix} \Sigma_{yy}(\theta) & \Sigma_{yx}(\theta) \\ \Sigma_{xy}(\theta) & \Sigma_{xx}(\theta) \end{bmatrix}$$

$$= \begin{bmatrix} \Lambda_y [(\mathbf{I} - \mathbf{B})^{-1} (\Gamma \Phi \Gamma' + \Psi)] [(\mathbf{I} - \mathbf{B})^{-1}]' \Lambda_y' + \Theta_\varepsilon & \Lambda_y (\mathbf{I} - \mathbf{B})^{-1} \Gamma \Phi \Lambda_x' \\ \Lambda_x \Phi \Gamma' [(\mathbf{I} - \mathbf{B})^{-1}]' \Lambda_y' & \Lambda_x \Phi \Lambda_x' + \Theta_\delta \end{bmatrix} \quad (1)$$

4.- Se autodescompone la matriz  $\Sigma(\theta)$  (Anexo B) a través de la descomposición de Cholesky (Raykov, Marcoulides y Boyd, 2003; Marcoulides, 2006), obteniendo la matriz triangular superior  $\mathbf{U}$ .

$$\text{Chol}(\Sigma(\theta)) = \mathbf{U} \cdot \mathbf{U}' \quad (2)$$

5.- Se genera una matriz  $\mathbf{Z}$  de datos aleatorios continuos con distribución normal, con tantas columnas como variables manifiestas y tantas filas como sujetos se haya establecido por condición. La cual se multiplica por la matriz  $\mathbf{U}$ , para obtener una matriz de datos muestrales  $\mathbf{X}$  con estructura conocida.

$$\mathbf{X} = \mathbf{Z} \cdot \mathbf{U} \quad (3)$$

6.- Se categoriza la matriz  $\mathbf{X}$  en 5 intervalos, de acuerdo con los umbrales presentados en la Tabla 3 (extraída y modificada de Garrido, Abad y Ponsoda, 2011), para la obtención de valores tipo Likert de 5 puntos, con distintos valores de asimetría en las variables manifiestas.

Tabla 3. Umbrales utilizados para obtener variables categóricas de 5 puntos, con tres niveles de asimetría.

Asimetría	Umbrales			
	1	2	3	4
0	-1.80	-0.60	0.60	1.80
1.25	0.24	0.65	1.06	1.54
2	0.68	1.00	1.34	1.77

Cada una de las muestras es sometida a las estimaciones pertinentes para recuperar los parámetros teóricos del modelo. De las estimaciones obtenidas, se valora el porcentaje de soluciones que convergen y el porcentaje de soluciones impropias, para cada una de las condiciones experimentales consideradas y para el total de muestras generadas. Se determinó un máximo de 100 iteraciones para lograr la convergencia de las soluciones y se consideró que una solución no es válida, cuando no converge o cuando la solución que converge incluye una o más varianzas negativas o incluye uno o más parámetros estimados estandarizados con valor superior a 1 en valor absoluto; de acuerdo con lo planteado por Lévy, Martín y Román (2006) y con el chequeo aproximado (*“rough check”*) señalado por Bollen (1989). Las muestras que presentan soluciones no válidas son excluidas de los análisis posteriores del estudio.

Para valorar la eficiencia en la recuperación de los parámetros se utiliza la varianza promedio ( $\overline{S^2}(\hat{\theta}_{jk})$ ) de las estimaciones de los J parámetros del Modelo Estructural, del Modelo de Medida Reflectivo y del Modelo de Medida Formativo, a través de las I réplicas válidas de cada condición K. Para ello, se calcula el promedio de las diferencias al cuadrado ( $\overline{d^2}$ ) entre el parámetro estimado ( $\hat{\theta}_{ijk}$ ) y la media de las estimaciones del parámetro j ( $\bar{\theta}_{jk}$ ) en las I réplicas válidas de cada condición intergrupo, para el Modelo Estructural, para el Modelo de Medida Reflectivo y para el Modelo de Medida Formativo:

$$\overline{d^2}_{jk} = \frac{\sum_{i=1}^I (\hat{\theta}_{ijk} - \bar{\theta}_{jk})^2}{I} \quad (4)$$



Luego, a través de un ANOVA de medidas repetidas se comparan las medias de las  $\bar{d}_{jk}^2$ , a través de las réplicas válidas, en cada condición experimental, correspondientes a las varianzas promedio del Modelo Estructural, a las varianzas promedio del Modelo de Medida Reflectivo y a las varianzas promedio del Modelo de Medida Formativo:

$$\bar{S}^2(\hat{\theta}_{ik}) = \frac{\sum_{j=1}^J (\bar{d}_{jk}^2)}{J} \quad (5)$$

Valores menores de varianza promedio de los parámetros estimados implican una mayor eficiencia en la estimación de los parámetros.

Para valorar el sesgo de las estimaciones, primero se calcula el promedio de las diferencias ( $\bar{d}_{jk}$ ) entre el parámetro estimado y el parámetro poblacional ( $\theta_{jk}$ ) para el Modelo Estructural, para el Modelo de Medida Reflectivo y para el Modelo de Medida Formativo:

$$\bar{d}_{jk} = \frac{\sum_{i=1}^I (\hat{\theta}_{ijk} - \theta_{jk})}{I} \quad (6)$$

donde  $\theta_{jk}$  es el valor del parámetro poblacional j en la condición k.

Luego, a través de un ANOVA de medidas repetidas se comparan las medias del sesgo promedio de los parámetros del Modelo Estructural, del Modelo de Medida Reflectivo y del Modelo de Medida Formativo, en las réplicas válidas por condición experimental, de acuerdo con:

$$\text{Sesgo}(\theta_{ik}) = \frac{\sum_{j=1}^J (\bar{d}_{jk})}{J} \quad (7)$$

Valores de sesgo inferiores al 10% son considerados aceptables (Bollen, et al., 2007; Muthén, du Toit y Spisic, 1997; Muthén, Kaplan y Hollis, 1987). Así, se consideran aceptables valores de sesgo inferiores a 0,06 para el Modelo Estructural, valores inferiores a 0,07 para el Modelo de Medida Reflectivo y valores inferiores a 0,04 para el Modelo de Medida Formativo.

Se calculó el Error Cuadrático Medio (ECM), por ser una medida habitualmente informada en estudios de recuperación de parámetros en MEE (p.e. Asparouhou y Muthén, 2009; Boomsma y Hoogland, 2001; Hwang y cols., 2010a; Ringle, Götz, Wetzels y Wilson, 2009) y utilizada para elegir entre estimadores, ya que considera la variabilidad de los estimadores y el sesgo de los mismos en una escala equivalente (la varianza de los estimadores y el sesgo cuadrado). Para ello, primero se calcula el promedio de las diferencias al cuadrado ( $\bar{d}_{jk}^2$ ) entre el parámetro estimado y el parámetro poblacional ( $\theta_j$ ) para el Modelo Estructural y para el Modelo de Medida, de acuerdo con:

$$\bar{d}_{jk}^2 = \frac{\sum_{i=1}^I (\hat{\theta}_{ijk} - \theta_{jk})^2}{I} \quad (8)$$

Luego, a través de un ANOVA de medidas repetidas se comparan las medias del ECM promedio del Modelo Estructural y del Modelo de Medida, en las réplicas válidas por condición experimental, de acuerdo con:

$$ECM(\hat{\theta}_{ik}) = \frac{\sum_{j=1}^J (\bar{d}_{jk}^2)}{J} \quad (9)$$

Valores menores de ECM implican una menor distancia entre los parámetros estimados y los parámetros poblacionales.

Siguiendo el planteamiento de Willmott y Matsuura (2005), quienes sugieren evitar utilizar medidas de error basadas en la suma de errores al cuadrado (p.e. RMSE), debido a la influencia que tienen sobre ellas la variabilidad de los errores. Para valorar la distancia entre los parámetros estimados y los parámetros poblacionales, se calculó la Media de las Diferencias Absolutas (MAD) entre el parámetro estimado y el parámetro poblacional, para el Modelo Estructural y del Modelo de Medida, de acuerdo con:

$$MAD(\hat{\theta}_{ik}) = \frac{\sum_{j=1}^J |\hat{\theta}_{ijk} - \theta_{jk}|}{J} \quad (10)$$

Luego, a través de un ANOVA de medidas repetidas se comparan las medias del MAD del Modelo Estructural y del Modelo de Medida, en las réplicas válidas por condición experimental. Valores menores de MAD implican una menor distancia entre la estimación del parámetro y el valor del parámetro poblacional.

## Análisis

Para la estimación de parámetros con ACOV, primero se calcula la matriz de correlaciones policóricas con el programa PRELIS (Anexo C), la cual se utiliza para la estimación de parámetros con el programa LISREL v.8.8 (Jöreskog y Sörbom, 2006;

Anexo D) utilizando como método de estimación Mínimos Cuadrados No Ponderados, siguiendo las directrices de Forero, Maydeu-Olivares y Gallardo-Pujol (2009).

Para el análisis con PLS-PM (Anexo E) se utiliza el programa R v. 2.14.0 (R Development core team, 2011) con el paquete “plsrm” (Sánchez y Trinchera, 2012). Considerando que estudios previos (Tenenhaus, Esposito Vinzi, Chatelin y Lauro, 2005; Ringle, Götz, Wetzels, Wilson, 2009) han comprobado la ausencia de diferencias en las estimación de parámetros al utilizar los distintos esquemas para la estimación de parámetros de PLS-PM, y el planteamiento de Esposito Vinzi, Trinchera y Amato (2010) se opta por utilizar el esquema de rutas, al ser el único esquema que considera la dirección de las relaciones como han sido establecidas en el modelo de rutas predictivo.

Para la automatización de los análisis en todas las condiciones se utiliza el programa Visual FoxPro 9.0 a través del cual se ejecutan las sentencias de análisis de LISREL (Anexo F) y R (Anexo G), y además se extraen las estimaciones estandarizadas de los parámetros desde los archivos de resultados de cada programa (Anexo H y Anexo I respectivamente).

Se realizan los 24000 análisis de estimación (6000 muestras  $\times$  2 técnicas de MEE  $\times$  2 especificaciones del modelo). Los datos resultantes se analizaron utilizando el programa IBM SPSS Statistics v.15. El modelo de análisis utilizado fue un Análisis de Varianzas (ANOVA) de medidas repetidas con 2 factores intra-grupos (técnica de MEE y especificación del modelo) y con 2 factores inter-grupos (nivel de asimetría y tamaño de la muestra), utilizando como variables dependientes la varianza de los errores de las estimaciones de los parámetros, el sesgo en las estimaciones de los parámetros, el error cuadrático medio y la media de las diferencias absolutas. Se valora el promedio por condición de los estadísticos mencionados de forma separada

para los parámetros del Modelo Estructural y para los parámetros del Modelo de Medida.

Para evaluar la relevancia estadística de los resultados se desestimó la utilización del valor de la significación estadística observada (valor p), debido a la gran cantidad de réplicas utilizadas y la influencia que tiene sobre el valor p el tamaño de la muestra. Por dichos motivos y al igual que en los estudios de simulación de Forero, Maydeu-Olivares y Gallardo-Pujol (2009), Hwang y cols. (2010), se optó por considerar el tamaño del efecto como valor de comparación, teniendo en cuenta sólo los resultados con tamaño del efecto al menos mediano ( $\eta^2 \geq 0,06$ ), de acuerdo con Cohen (1988). Se informa el valor de Eta cuadrado que asume esfericidad, considerando que éste, es el mismo valor que se obtuvo para Greenhouse-Geisser, Huynh-Feldt y Límite-inferior.

## **Resultados**

El porcentaje de estimaciones convergentes en el total de réplicas fue elevado tanto para ACOV (97,35%) como para PLS-PM (98,24%). El menor porcentaje de convergencia por condición fue de 82,6% para ACOV con el modelo menos especificado, asimetría de 2 y tamaño muestral de 50; condición en la que PLS-PM presentó un 94% de soluciones que convergieron (Tabla 4). Se desestimó la utilización del 4% del total de réplicas que no convergen tras 100 iteraciones en los análisis posteriores.

Tabla 4. Tabla de frecuencias de réplicas que no convergen tras 100 iteraciones\*, con ACOV y PLS-PM.

Asimetría	n	ACOV			PLS-PM		
		Modelo Más especificado	Modelo Menos especificado	Total	Modelo Más especificado	Modelo Menos especificado	Total
0	50	1(99.8)	9(98.2)	9(98.2)	0(100)	6(98.8)	6(98.8)
	100	1(99.8)	1(99.8)	2(99.6)	0(100)	1(99.8)	1(99.8)
	300	0(100)	0(100)	0(100)	0(100)	0(100)	0(100)
	500	0(100)	0(100)	0(100)	0(100)	0(100)	0(100)
1.25	50	11(97.8)	18(96.4)	24(95.2)	7(98.6)	10(98)	15(97)
	100	0(100)	2(99.6)	2(99.6)	0(100)	3(99.4)	3(99.4)
	300	0(100)	0(100)	0(100)	0(100)	0(100)	0(100)
	500	1(99.8)	0(100)	1(99.8)	0(100)	0(100)	0(100)
2	50	72(85.6)	87(82.6)	104(79.2)	47(90.6)	30(94)	66(86.8)
	100	5(99)	13(97.4)	13(97.4)	9(98.2)	7(98.6)	15(97)
	300	2(99.6)	1(99.8)	3(99.4)	0(100)	0(100)	0(100)
	500	0(100)	1(99.8)	1(99.8)	0(100)	0(100)	0(100)
Total		93(98.5)	132(97.8)	159(97.4)	63(99)	57(99.1)	106(98.2)

\* Entre paréntesis se presenta el porcentaje de réplicas que convergen.

En cuanto a la presencia de soluciones impropias, PLS-PM presentó un 0,2% de soluciones impropias, en cambio ACOV presentó un 15,2% de soluciones impropias del total de réplicas generadas. El mayor porcentaje de soluciones impropias (62,2%) lo presentó ACOV en la condición de modelo menos especificado, con asimetría de 2 y un tamaño muestral de 50, instancia en que PLS-PM presentó un 0,8% de soluciones impropias.

Las réplicas que presentaron soluciones impropias no fueron consideradas en los análisis posteriores. En total se desestimó la utilización de un 17,70% de las réplicas, considerando las 4938 réplicas válidas para los análisis posteriores.

### *Eficiencia*

En el ANOVA de medidas repetidas (2x2x3x4) que valoraba la varianza promedio de las estimaciones estandarizadas de los parámetros del Modelo Estructural (Tabla 5.A), no se evidenciaron interacciones de tercer o segundo orden con tamaños del efecto medianos o superiores. Se encontraron los siguientes efectos de interacción de primer orden:

Tabla 5. Pruebas de efectos intra-sujetos e inter-grupo para la varianza promedio de las estimaciones estandarizadas de los parámetros del Modelo Estructural (A) y del Modelo de Medida (B).

Fuente	gl	(A) Modelo Estructural			(B) Modelo de Medida		
		F	Sig	$\eta^2$	F	Sig	$\eta^2$
Técnica de MEE (Técnica)	1	1527.6	.00	.24**	271.2	.00	.05
Especificación del modelo (Espec)	1	543.4	.00	.10**	1.8	.18	.00
Nivel de Asimetría (Asim)	2	753.2	.00	.23**	624.4	.00	.20**
Tamaño muestral (n)	3	33.0	.00	.02	44.3	.00	.03
Técnica*Espec	1	767.3	.00	.09**	.1	.73	.00
Técnica*Asim	2	387.1	.00	.14**	35.2	.00	.01
Técnica*n	3	17.8	.00	.01	11.5	.00	.01
Espec*Asim	2	179.9	.00	.07**	.3	.78	.00
Espec*n	3	16.9	.00	.01	.3	.85	.00
Asim * n	6	44.3	.00	.05	61.1	.00	.07**
Técnica*Espec*Asim	2	106.0	.00	.04	.1	.89	.00
Técnica*Espec*n	3	4.6	.00	.00	.2	.92	.00
Técnica*Asim*n	6	25.0	.00	.03	10.0	.00	.01
Espec*Asim*n	6	25.1	.00	.03	.9	.49	.00
Técnica*Espec*Asim*n	6	10.3	.00	.01	.6	.75	.00

\*\* Tamaño del efecto al menos mediano ( $\eta^2 \geq 0,06$ ).

Técnica de MEE y nivel de asimetría ( $\eta^2=0,14$ ). PLS-PM presentó menores valores de varianza promedio que ACOV, en asimetría nula (Figura 7). La diferencia fue menor cuando se utilizó asimetría media y los valores de varianza promedio fueron similares para ambas técnicas de MEE cuando se utilizó un nivel de asimetría alta.

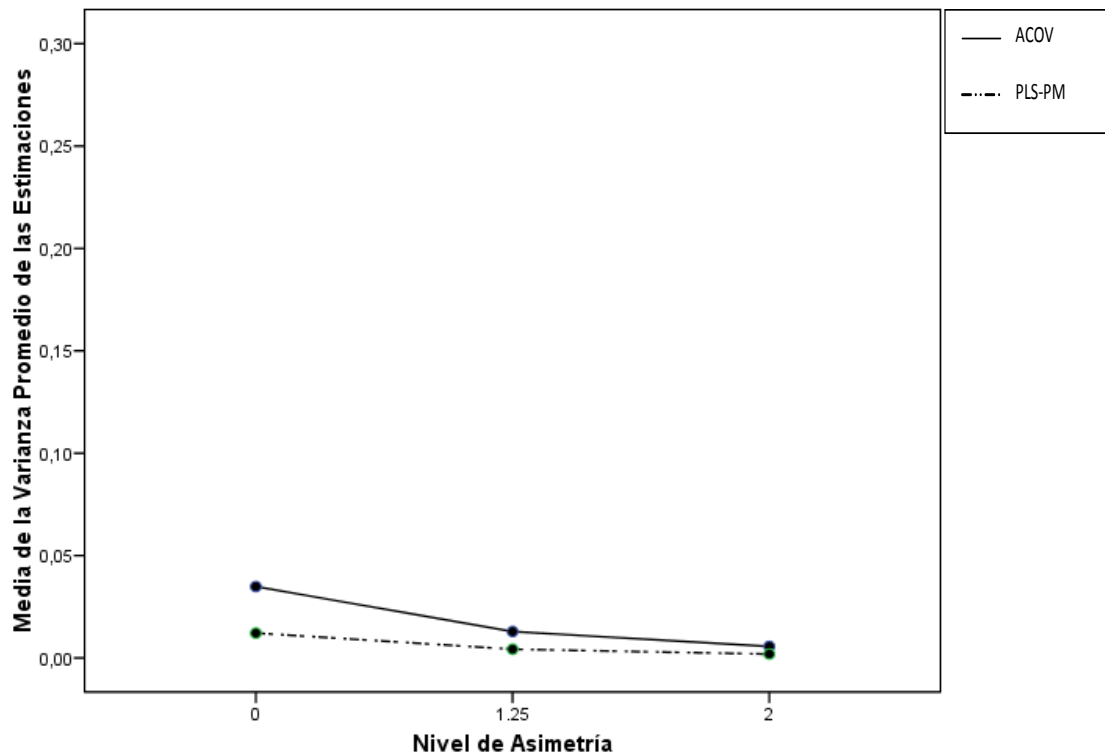


Figura 7. Medias de la varianza promedio de las estimaciones del Modelo Estructural con ACOV y PLS-PM, de acuerdo al nivel de asimetría.

Técnica de MEE y especificación del modelo ( $\eta^2=0,09$ ). Las estimaciones de PLS-PM presentaron una menor varianza promedio de las estimaciones, cuando se utilizó el modelo menos especificado. Cuando se utilizó el modelo más especificado ambas técnicas de MEE obtuvieron varianzas promedio de las estimación similares (Figura 8).



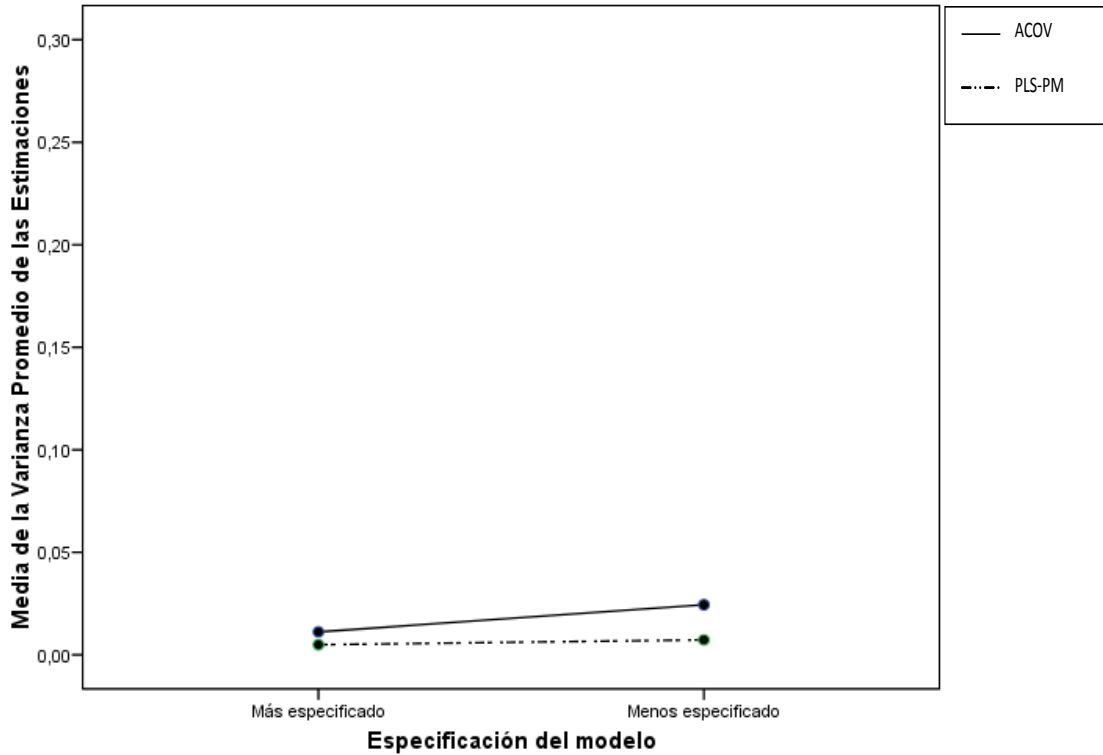


Figura 8. Medias de la varianza promedio de las estimaciones del Modelo Estructural con ACOV y PLS-PM, de acuerdo a la especificación del modelo.

Especificación del modelo y nivel de asimetría ( $\eta^2=0,07$ ). Las estimaciones con el modelo más especificado presentaron una menor varianza promedio de las estimaciones cuando se utilizó un nivel de asimetría nula. La diferencia fue menor cuando se utilizó un nivel de asimetría medio. Y la varianza promedio con ambas modalidades de especificación del modelo fue similar cuando se utilizó un nivel de asimetría alto.

En el ANOVA de medidas repetidas ( $2 \times 2 \times 3 \times 4$ ) para la varianza promedio de las estimaciones estandarizadas de los parámetros del Modelo de Medida (Tabla 5.b), no se evidenciaron interacciones de tercer o segundo orden. Se encontró una interacción de primer orden entre el nivel de asimetría y el tamaño muestral ( $\eta^2=0,07$ ). La varianza promedio de las estimaciones cuando se utilizó asimetría nula tendió a ser

menor en los distintos tamaños muestrales. La varianza promedio de las estimaciones cuando se utilizó asimetría media fue similar a la varianza promedio de las estimaciones cuando se utilizó asimetría nula. La varianza promedio de las estimaciones cuando se utilizó asimetría alta fue mayor que la varianza promedio de las estimaciones con los otros dos niveles de asimetría, particularmente con un tamaño muestral igual a 100.

### Sesgo

En el ANOVA de medidas repetidas (2x2x3x4) para el sesgo promedio de las estimaciones de los parámetros del Modelo Estructural (Tabla 6.A), no se evidenciaron interacciones de tercer o segundo orden con tamaños del efecto medianos o superiores. Se encontraron los siguientes efectos de interacción de primer orden:

Tabla 6. Pruebas de efectos intra-sujetos e inter-sujetos para el sesgo promedio de las estimaciones estandarizadas de los parámetros del Modelo Estructural (A) y del Modelo de Medida (B).

Fuente	(A) Modelo Estructural				(B) Modelo de Medida		
	gl	F	Sig	$\eta^2$	F	Sig	$\eta^2$
Técnica de MEE (Técnica)	1	88802.5	.00	.95**	74665.7	.00	.94**
Especificación del modelo (Espec)	1	16542.7	.00	.76**	262.8	.00	.05
Nivel de asimetría (Asim)	2	45.6	.00	.02	274.7	.00	.10**
Tamaño muestral (n)	3	55.1	.00	.03	14.3	.00	.01
Técnica*Espec	1	18792.3	.00	.79**	101.9	.00	.02
Técnica*Asim	2	833.7	.00	.25**	2541.1	.00	.50**
Técnica*n	3	39.8	.00	.02	174.5	.00	.08**
Espec*Asim	2	48.9	.00	.02	41.3	.00	.02
Espec*n	3	89.9	.00	.05	81.6	.00	.05
Asim * n	6	9.5	.00	.01	2.3	.03	.00
Técnica*Espec*Asim	2	135.4	.00	.05	2.5	.08	.00
Técnica*Espec*n	3	8.9	.00	.01	12.7	.00	.01
Técnica*Asim*n	6	6.0	.00	.01	10.7	.00	.01
Espec*Asim*n	6	12.8	.00	.02	13.4	.00	.02
Técnica*Espec*Asim*n	6	0.9	.53	.00	3.6	.00	.00

\*\* Tamaño del efecto al menos mediano ( $\eta^2 \geq 0,06$ ).

Técnica de MEE y especificación del modelo ( $\eta^2=0,79$ ). El sesgo promedio de las estimaciones del Modelo Estructural de ACOV, fue levemente superior a 0, encontrándose dentro del rango aceptable de sesgo, en las dos modalidades de especificación del modelo (Figura 9). En cambio las estimaciones de PLS-PM presentaron un sesgo promedio negativo que se encontraba en un rango inaceptable en ambas modalidades de especificación del modelo. El sesgo promedio de las estimaciones con PLS-PM tendió a disminuir en condiciones de modelo menos especificado.

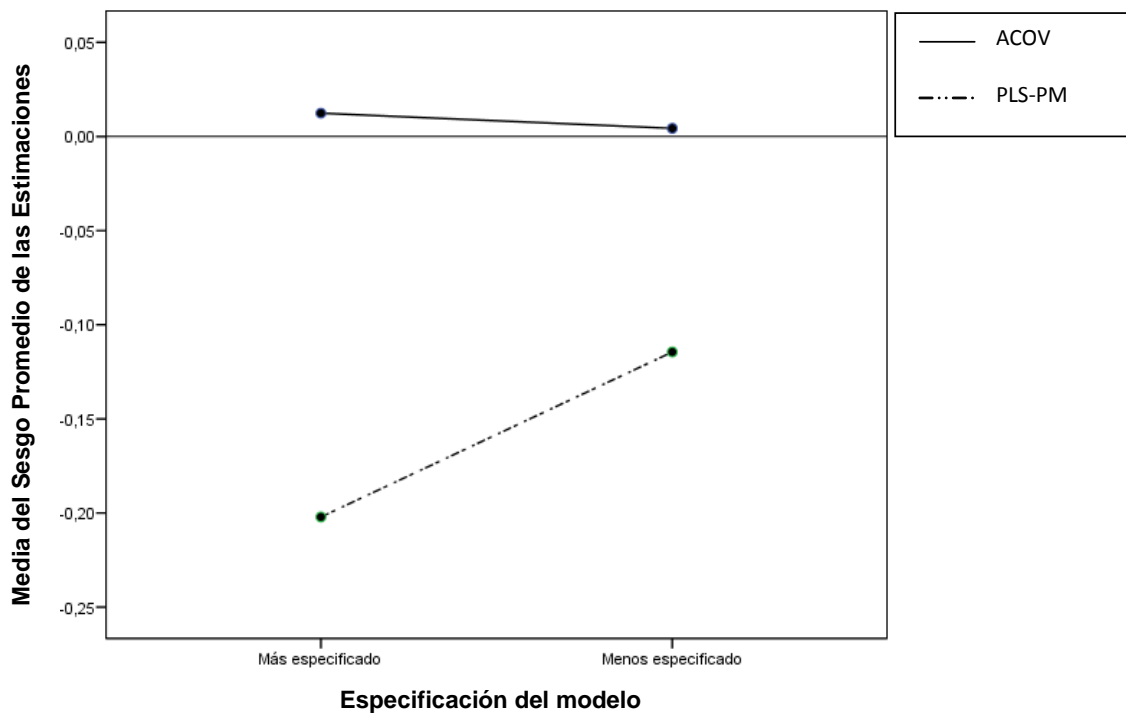


Figura 9. Medias del sesgo promedio de las estimaciones del Modelo Estructural con ACOV y PLS-PM, de acuerdo a la especificación del modelo.

Técnica de MEE y nivel de asimetría ( $\eta^2=0,25$ ). El sesgo promedio de la estimación de parámetros del Modelo Estructural de ACOV, fue levemente superior a 0, encontrándose dentro del rango aceptable en los tres niveles de asimetría, siendo levemente mayor para una asimetría de 2 (Figura 10). El sesgo promedio de las

estimaciones de PLS-PM fue negativo, se encontró fuera del rango aceptable y aumentó en la medida que aumentaba el nivel de asimetría.

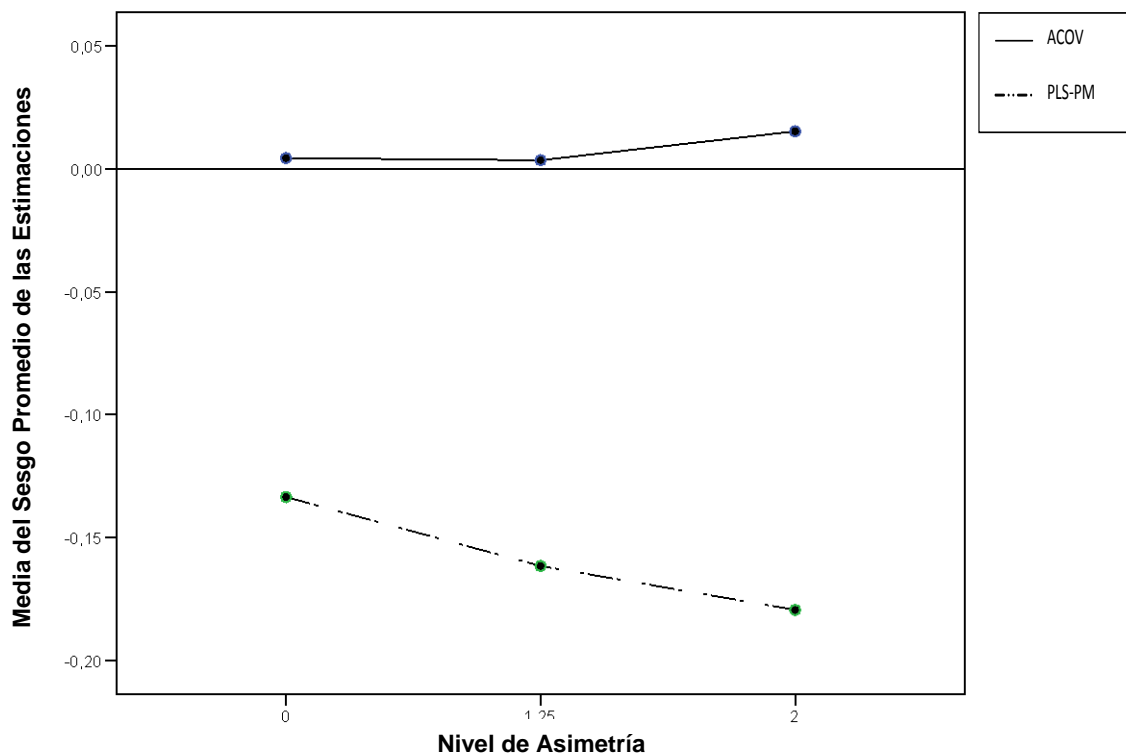


Figura 10. Medias del sesgo promedio de las estimaciones del Modelo Estructural con ACOV y PLS-PM, de acuerdo al nivel de asimetría.

En el ANOVA de medidas repetidas (2x2x3x4) para el sesgo promedio de las estimaciones de los parámetros del Modelo de Medida (Tabla 6.B), no se evidenciaron interacciones de tercer o segundo orden con tamaños del efecto medianos o superiores. Se encontraron los siguientes efectos de interacción de primer orden:

Técnica de MEE y nivel de asimetría ( $\eta^2=0,50$ ). El sesgo promedio de las estimaciones de parámetros del Modelo de Medida de ACOV, fue levemente inferior a 0, encontrándose dentro del rango aceptable en los tres niveles de asimetría (Figura

11). En cambio el sesgo promedio de las estimaciones de PLS-PM fue positivo y se encontró fuera del rango aceptable para asimetría igual a 0, disminuyendo dentro del rango normal para niveles de asimetría de 1,25 y disminuyendo aún más con una asimetría de 2.

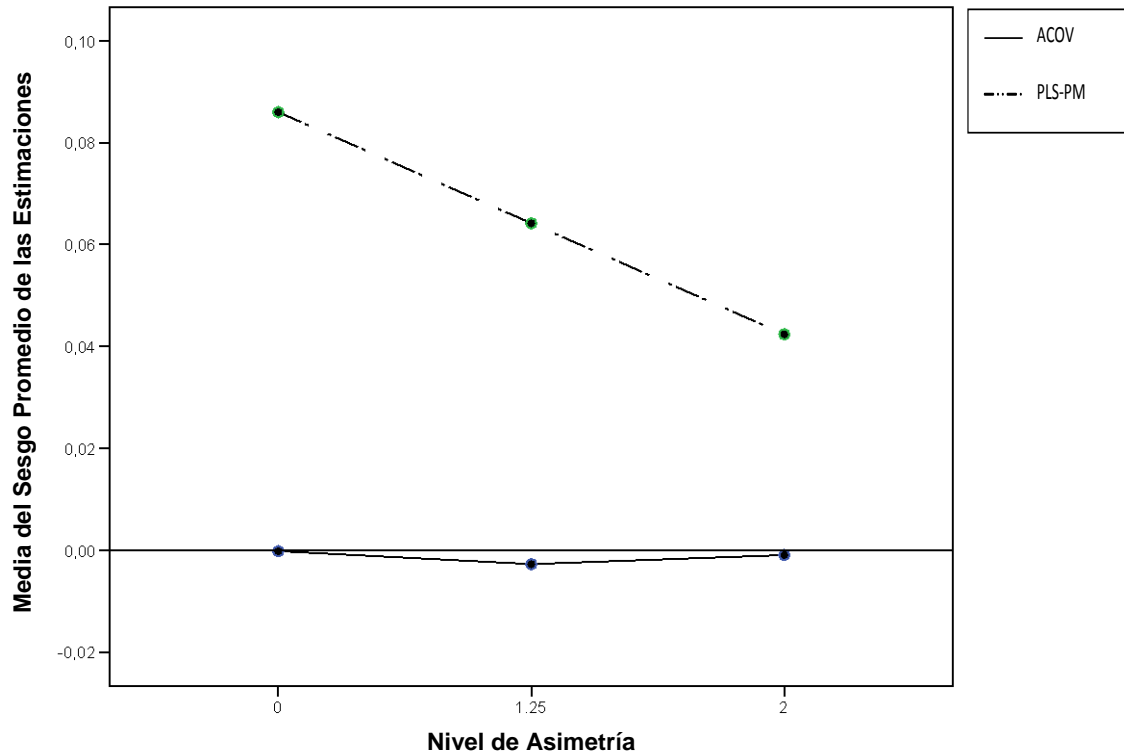


Figura 11. Medias del sesgo promedio de las estimaciones del Modelo de Medida con ACOV y PLS-PM, de acuerdo al nivel de asimetría.

Técnica de MEE y tamaño de la muestra ( $\eta^2=0,08$ ). El sesgo promedio de las estimaciones de parámetros del Modelo de Medida de ACOV fue levemente inferior a 0, encontrándose dentro del rango aceptable en los distintos tamaños muestrales (Figura 12). Por su parte las estimaciones de PLS-PM presentaron un sesgo promedio positivo, el cual, aunque se encontró dentro del rango normal, fue muy próximo al límite superior y tendió a aumentar cuando aumentaba el tamaño muestral.

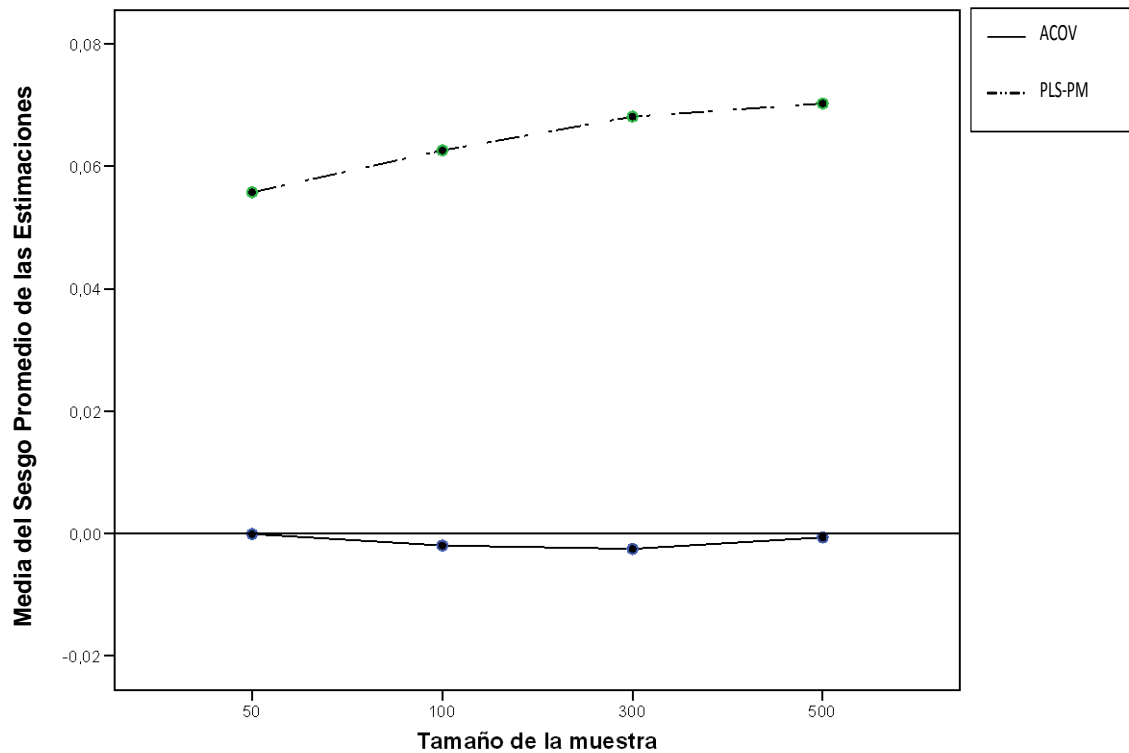


Figura 12. Medias del sesgo promedio de las estimaciones del Modelo de Medida con ACOV y PLS-PM, de acuerdo al tamaño de la muestra.

### *Error Cuadrático Medio (ECM)*

En el ANOVA de medidas repetidas (2x2x3x4) que valoraba el ECM promedio de las estimaciones estandarizadas de los parámetros del Modelo Estructural (Tabla 7.A), no se evidenciaron interacciones de tercer o segundo orden con tamaños del efecto medianos o superiores. Se encontraron los siguientes efectos de interacción de primer orden:

Tabla 7. Pruebas de efectos intra-sujetos e inter-grupo para el ECM promedio de las estimaciones estandarizadas de los parámetros del Modelo Estructural (A) y del Modelo de Medida (B).

Fuente	(A) Modelo Estructural				(B) Modelo de Medida		
	gl	F	Sig	$\eta^2$	F	Sig	$\eta^2$
Técnica de MEE (Técnica)	1	4221.9	.00	.46**	185.2	.00	.04
Especificación del modelo (Espec)	1	201.2	.00	.04	9.3	.00	.00
Nivel de Asimetría (Asim)	2	245.7	.00	.09**	724.7	.00	.23**
Tamaño muestral (n)	3	45.8	.00	.03	70.9	.00	.04
Técnica*Espec	1	705.1	.00	.13**	2.6	.11	.00
Técnica*Asim	2	476.1	.00	.16**	85.9	.00	.03
Técnica*n	3	185.6	.00	.10**	30.2	.00	.02
Espec*Asim	2	196.4	.00	.07**	2.0	.14	.00
Espec*n	3	27.0	.00	.02	1.6	.19	.00
Asim * n	6	60.3	.00	.07**	53.7	.00	.06**
Técnica*Espec*Asim	2	22.7	.00	.01	.2	.79	.00
Técnica*Espec*n	3	12.8	.00	.01	.8	.47	.00
Técnica*Asim*n	6	3.8	.00	.01	12.4	.00	.02
Espec*Asim*n	6	13.0	.00	.02	3.3	.00	.00
Técnica*Espec*Asim*n	6	9.0	.00	.01	2.4	.02	.00

\*\* Tamaño del efecto al menos mediano ( $\eta^2 \geq 0,06$ ).

Técnica de MEE y nivel de asimetría ( $\eta^2=0,16$ ). El ECM promedio de las estimaciones de ACOV fue menor que el ECM promedio de PLS-PM en los distintos niveles de asimetría utilizados. La diferencia fue menor cuando se utilizó asimetría nula y tendió a aumentar cuando se utilizó asimetría media y alta. En dichas condiciones el ECM promedio de ACOV disminuyó y el ECM promedio de PLS-PM se mantuvo en valores similares a los presentados en asimetría nula (Figura 13).

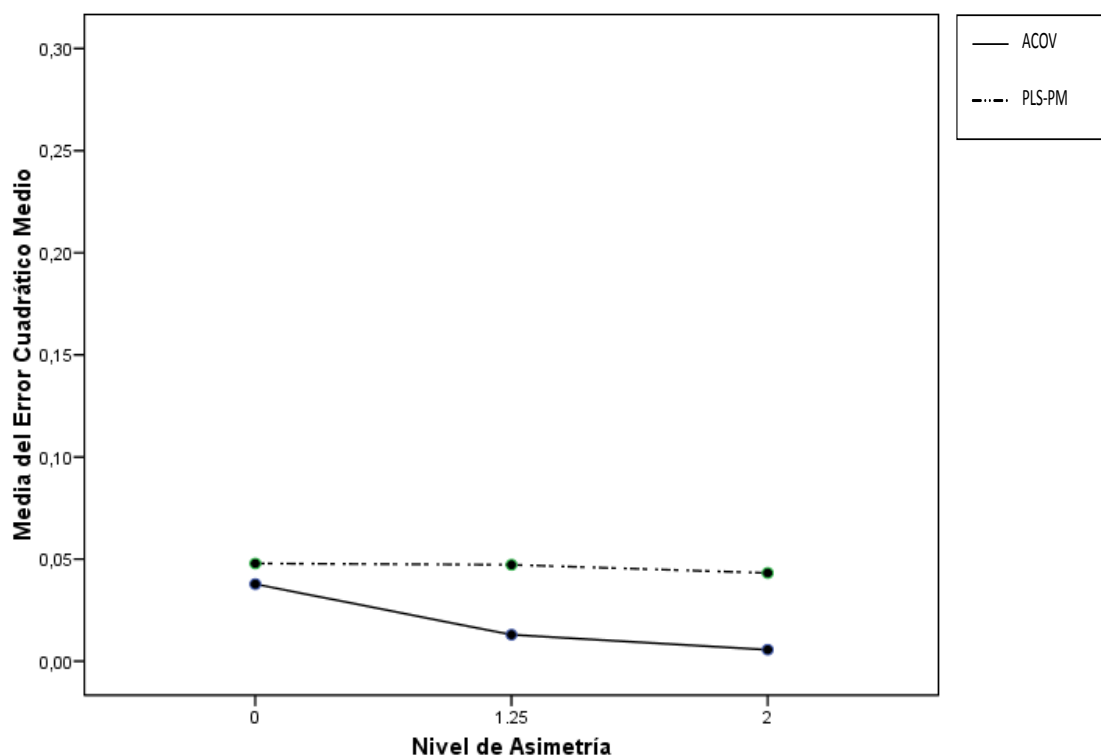


Figura 13. Medias del EMC de las estimaciones del Modelo Estructural con ACOV y PLS-PM, de acuerdo al nivel de asimetría.

Técnica de MEE y especificación del modelo ( $\eta^2=0,13$ ). El ECM promedio de las estimaciones de ACOV fue menor que el ECM de PLS-PM en las dos modalidades de especificación del modelo. La diferencia fue mayor cuando se utilizó el modelo más especificado. Cuando se utilizó el modelo menos especificado el ECM promedio de ACOV tendió a aumentar y el ECM promedio de PLS-PM se mantuvo en valores similares a los presentados con el modelo más especificado (Figura 14).



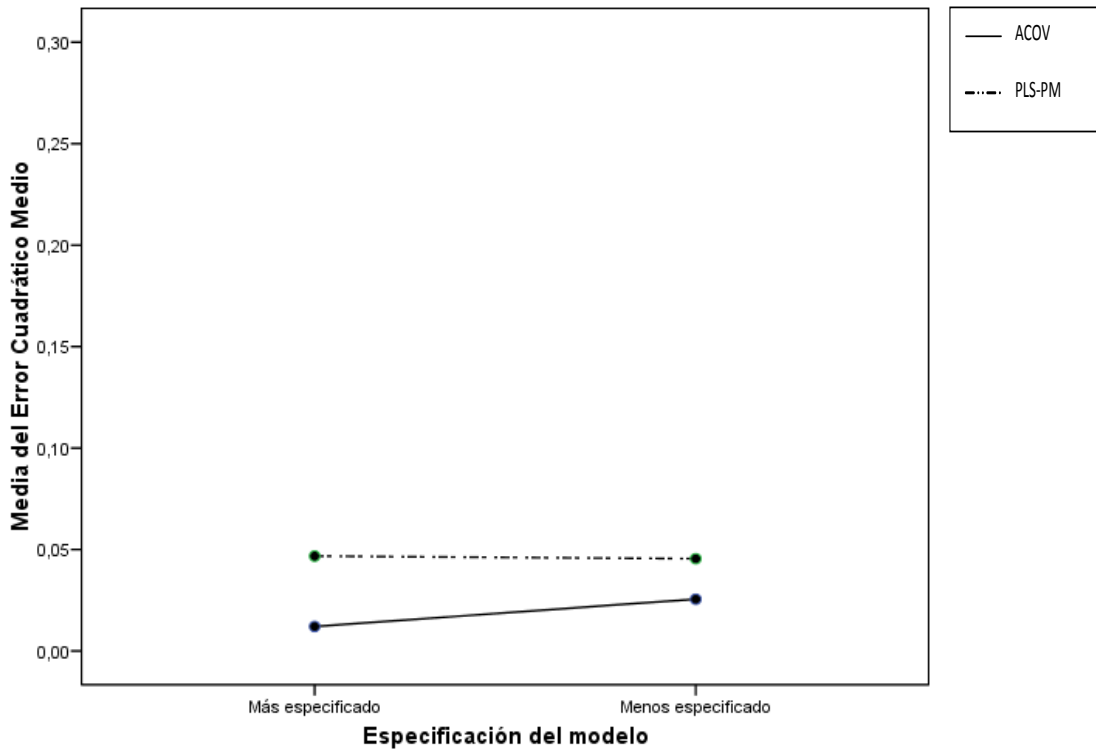


Figura 14. Medias del EMC de las estimaciones del Modelo Estructural con ACOV y PLS-PM, de acuerdo con la especificación del modelo.

Técnica de MEE y tamaño de la muestra ( $\eta^2=0,10$ ). El ECM promedio de las estimaciones de ACOV fue menor que el ECM promedio de PLS-PM en los distintos tamaños muestrales. La diferencia fue similar cuando se utilizó un tamaño muestral pequeño y mediano, y tendió a aumentar cuando se utilizó un tamaño muestral grande y muy grande. En dicha condiciones el ECM promedio de ACOV disminuyó y el ECM promedio de PLS-PM tendió a aumentar (Figura 15).

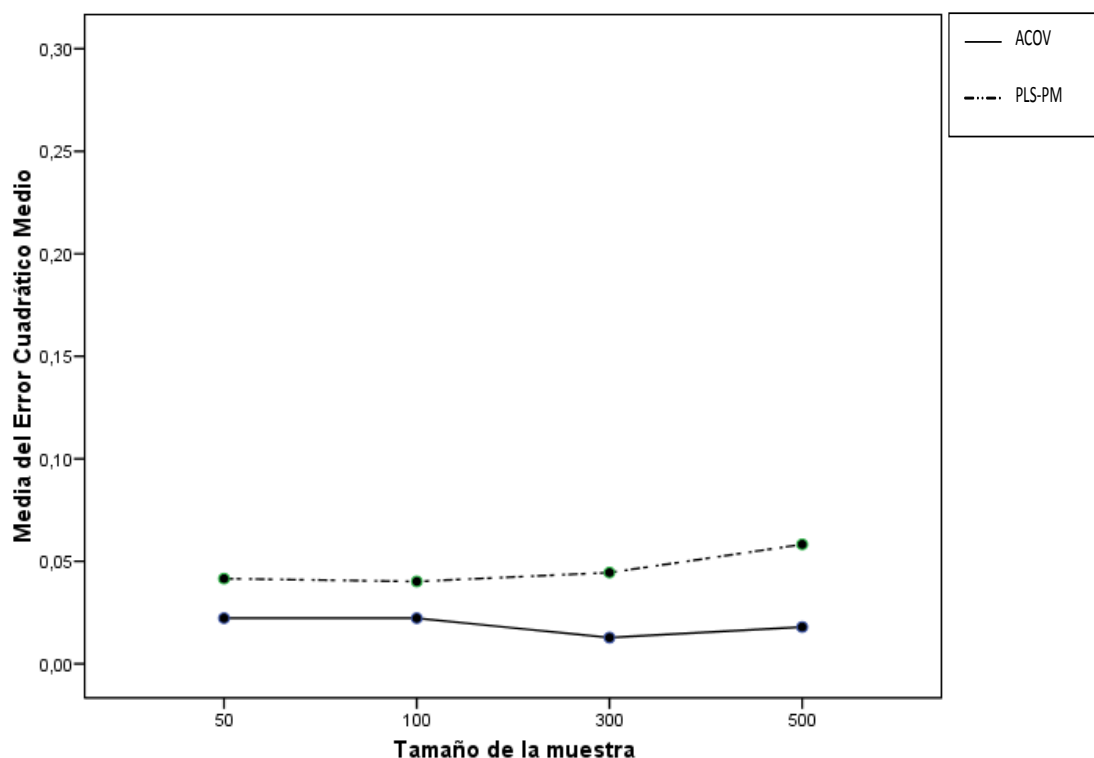


Figura 15. Medias del EMC de las estimaciones del Modelo Estructural con ACOV y PLS-PM, de acuerdo el tamaño muestral.

Especificación del modelo y nivel de asimetría ( $\eta^2=0,07$ ). Las estimaciones con el modelo más especificado presentaron un menor ECM promedio que las estimaciones con el modelo menos especificado cuando se utilizó asimetría nula. En niveles de asimetría medio y alto el ECM promedio fue similar para las dos modalidades de especificación del modelo.

Nivel de asimetría y tamaño de la muestra ( $\eta^2=0,07$ ). Las estimaciones con los tres niveles de asimetría presentaron un ECM promedio similar cuando se utilizó un tamaño muestral grande y muy grande. Cuando se utilizó un tamaño muestral mediano las estimaciones con asimetría nula presentaron un mayor ECM promedio que las estimaciones con los otros dos niveles de asimetría, los que tendieron a presentar

valores de ECM promedio similares. Cuando se utilizó un tamaño muestral pequeño el mayor ECM promedio lo obtuvieron las estimaciones con asimetría nula y el menor ECM promedio lo obtuvieron las estimaciones con asimetría alta.

En el ANOVA de medidas repetidas (2x2x3x4) que valoraba el ECM promedio de las estimaciones estandarizadas de los parámetros del Modelo de Medida (Tabla 7.B), no se evidenciaron interacciones de tercer o segundo orden con tamaños del efecto medianos o superiores. Se encontraron el siguiente efecto de interacción de primer orden:

Nivel de asimetría y tamaño de la muestra ( $\eta^2=0,06$ ). Las estimaciones con los tres niveles de asimetría presentaron un ECM promedio similar cuando se utilizó un tamaño muestral grande y muy grande. Cuando se utilizó un tamaño muestral mediano las estimaciones con asimetría nula presentaron un mayor ECM promedio que las estimaciones con los otros dos niveles de asimetría, los que presentaron valores similares. Cuando se utilizó un tamaño muestral pequeño el mayor ECM promedio lo obtuvieron las estimaciones con asimetría nula y el menor ECM promedio lo obtuvieron las estimaciones con asimetría alta, aunque las diferencias fueron pequeñas.

#### *Media de las Diferencias Absolutas (MAD)*

En el ANOVA de medidas repetidas (2x2x3x4) que valoraba el MAD de las estimaciones estandarizadas de los parámetros del Modelo Estructural (Tabla 8.A), no se evidenciaron interacciones de tercer orden con tamaños del efecto medianos o superiores. Se encontraron el siguiente efecto de interacción de segundo orden:

Tabla 8. Pruebas de efectos intra-sujetos e inter-grupo para el MAD de las estimaciones estandarizadas de los parámetros del Modelo Estructural (A) y del Modelo de Medida (B).

Fuente	(A) Modelo Estructural				(B) Modelo de Medida		
	gl	F	Sig	$\eta^2$	F	Sig	$\eta^2$
Técnica de MEE (Técnica)	1	8162.0	.00	.62**	2183.5	.00	.31**
Especificación del modelo (Espec)	1	217.9	.00	.04	71.5	.00	.01
Nivel de Asimetría (Asim)	2	243.2	.00	.09**	145.4	.00	.06**
Tamaño muestral (n)	3	325.5	.00	.17**	2828.8	.00	.63**
Técnica*Espec	1	1153.5	.00	.19**	16.9	.00	.00
Técnica*Asim	2	11.1	.00	.00	703.7	.00	.22**
Técnica*n	3	1180.3	.00	.42**	582.3	.00	.26**
Espec*Asim	2	132.4	.00	.05	18.4	.00	.01
Espec*n	3	62.3	.00	.04	12.4	.00	.01
Asim * n	6	6.9	.00	.01	94.0	.00	.10**
Técnica*Espec*Asim	2	216.8	.00	.08**	12.7	.00	.01
Técnica*Espec*n	3	30.6	.00	.02	6.8	.00	.00
Técnica*Asim*n	6	31.2	.00	.04	64.8	.00	.07**
Espec*Asim*n	6	9.4	.00	.01	9.3	.00	.01
Técnica*Espec*Asim*n	6	6.7	.00	.01	5.3	.00	.01

\*\* Tamaño del efecto al menos mediano ( $\eta^2 \geq 0,06$ ).

Técnica de MEE, especificación del modelo y nivel de asimetría ( $\eta^2=0,08$ ). Las estimaciones de ACOV presentaron un menor promedio de MAD, que las estimaciones de PLS-PM, en las dos modalidades de especificación del modelo, cuando se utilizaron los distintos niveles de asimetría. La diferencia entre el MAD de ACOV y PLS-PM fue menor con el modelo menos especificado y dicha diferencia disminuyó en la medida que aumentaba el nivel de asimetría. La diferencia entre el MAD de las técnicas con el modelo más especificado aumentó en la medida que aumentó el nivel de asimetría (Figuras 16, 17 y 18).

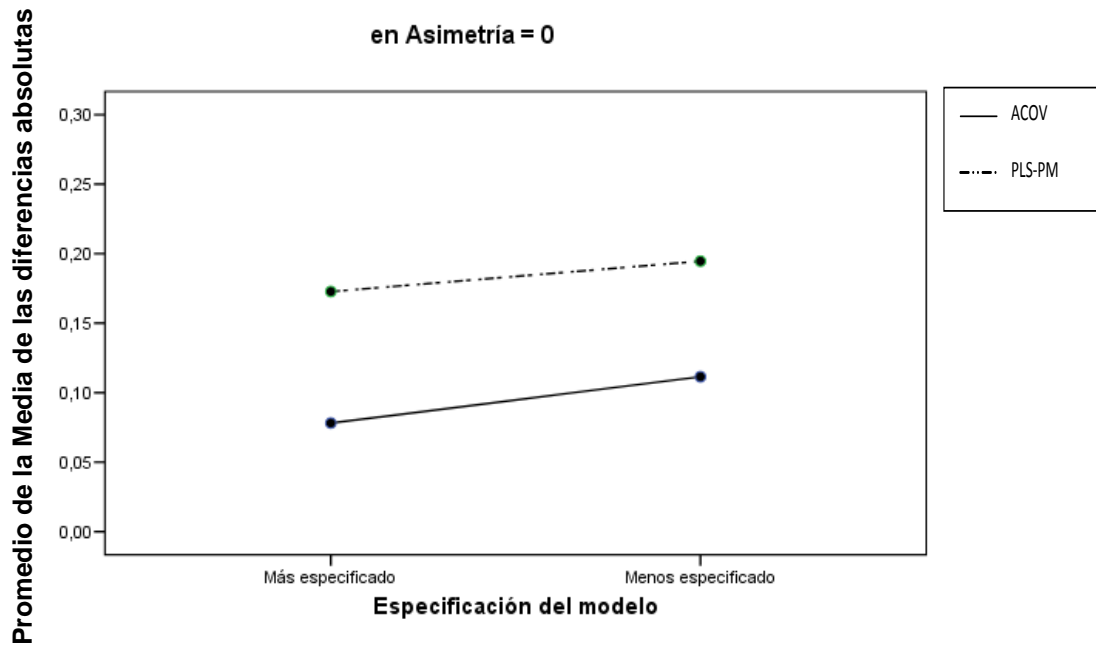


Figura 16. Promedios del MAD de las estimaciones del Modelo Estructural con ACOV y PLS-PM, de acuerdo a la especificación del modelo, con asimetría nula.

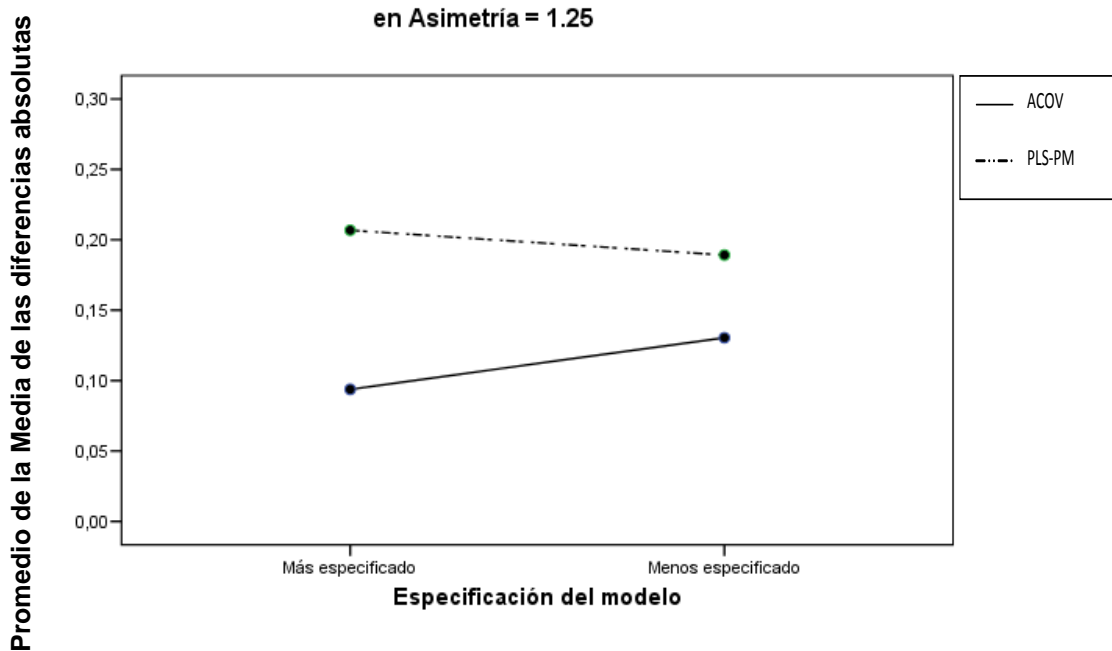


Figura 17. Promedios del MAD de las estimaciones del Modelo Estructural con ACOV y PLS-PM, de acuerdo a la especificación del modelo, con asimetría media.

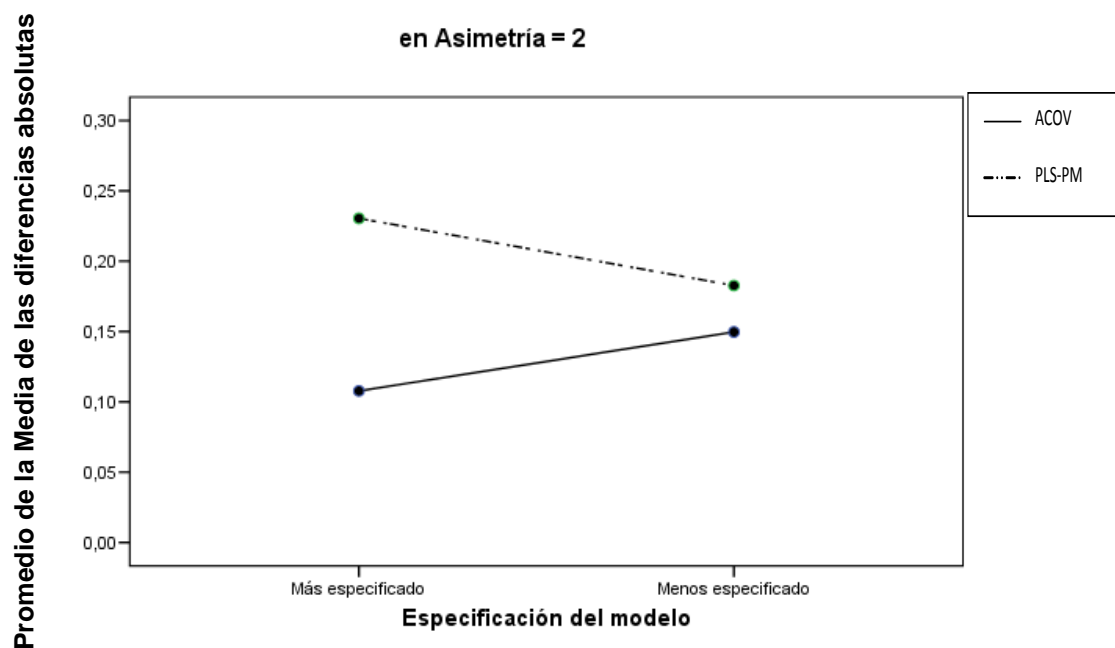


Figura 18. Promedios del MAD de las estimaciones del Modelo Estructural con ACOV y PLS-PM, de acuerdo a la especificación del modelo, con asimetría alta.

Además se encontró el efecto de interacción de primer orden entre la Técnica de MEE y tamaño de la muestra ( $\eta^2=0,42$ ). El promedio del MAD de las estimaciones de ACOV fue menor que el promedio del MAD de PLS-PM cuando se utilizaron tamaños muestrales muy grandes, grandes y medianos. El promedio de MAD de PLS-PM tendió a aumentar en la medida que disminuía el tamaño muestral, y el promedio del MAD de ACOV tendió a aumentar en la medida que disminuía el tamaño de la muestra, llegando a niveles similares a los presentados por PLS-PM cuando el tamaño de las muestras fue pequeño. Así, las diferencias entre los promedios de los MAD disminuyó en la medida que disminuyó el tamaño muestral (Figura 19).

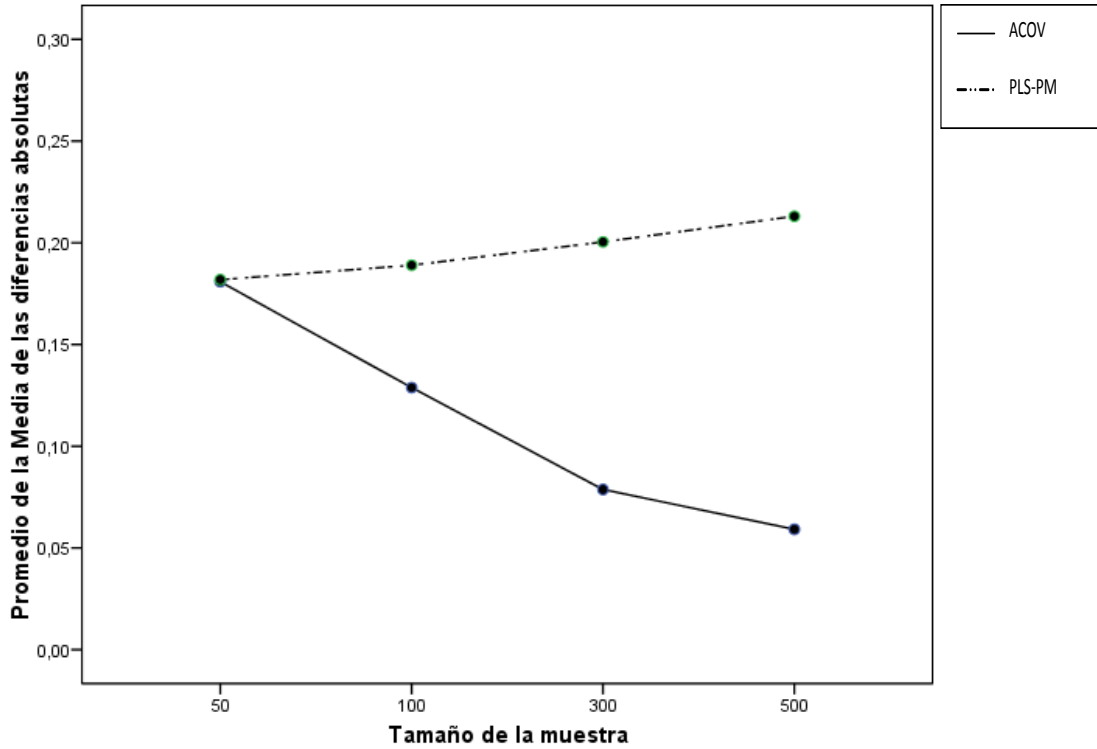


Figura 19. Promedios del MAD de las estimaciones del Modelo Estructural con ACOV y PLS-PM, de acuerdo al tamaño de la muestra.

En el ANOVA de medidas repetidas (2x2x3x4) para el MAD de las estimaciones estandarizadas de los parámetros del Modelo de Medida (Tabla 8.B), no se evidenció un tamaño del efecto mediano o superior para la interacción de tercer orden. Se encontró un efecto de interacción de segundo orden entre Técnica de MEE, nivel de asimetría y tamaño muestral ( $\eta^2=0,07$ ). Cuando se utilizó asimetría nula, el promedio del MAD para las estimaciones de ACOV fue menor, con muestras muy grandes y grandes. Estas diferencias disminuyeron en la medida que aumentaba la asimetría, llegando a valores muy similares cuando se utilizó asimetría alta. Con muestras medianas y pequeñas, el promedio del MAD fue similar para ambas técnicas en los distintos niveles de asimetría (Figura 20, 21 y 22).

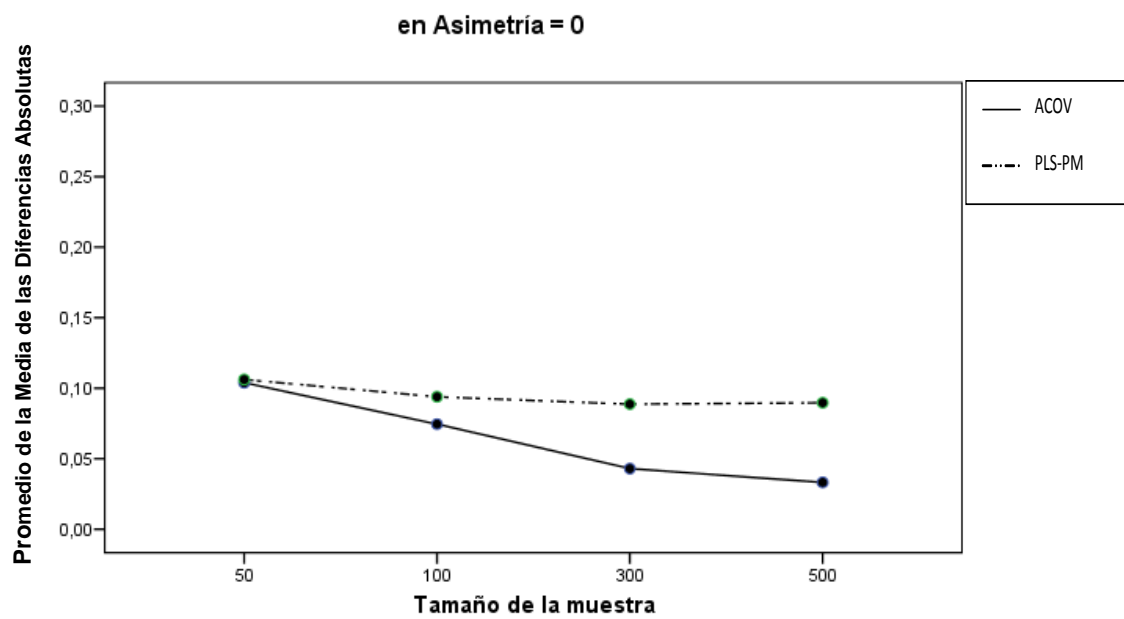


Figura 20. Promedios para el MAD de las estimaciones del Modelo de Medida con ACOV y PLS-PM, de acuerdo al tamaño de la muestra con asimetría nula.

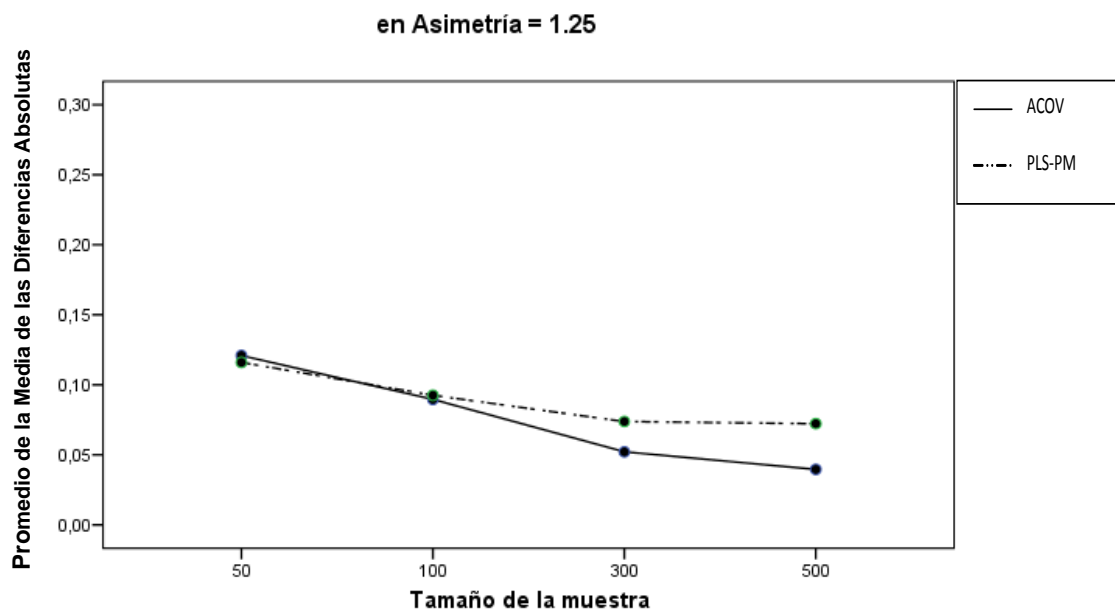


Figura 21. Promedio del MAD de las estimaciones del Modelo de Medida con ACOV y PLS-PM, de acuerdo al tamaño de la muestra con asimetría media.



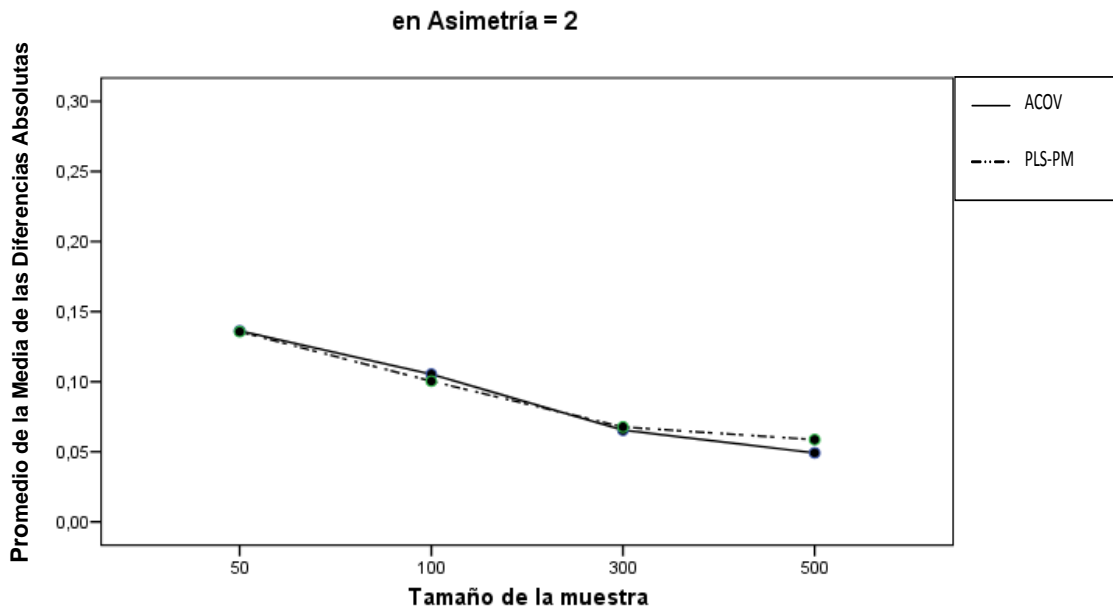


Figura 22. Promedio del MAD de las estimaciones del Modelo de Medida con ACOV y PLS-PM, de acuerdo al tamaño de la muestra con asimetría alta.

Por último, respecto a la incorporación del efecto impropio del parámetro estructural  $\gamma_{21}$ , PLS-PM presentó estimaciones superiores ( $\bar{x} = 0,11$ ;  $dt = 0,10$ ;  $s / \sqrt{n} = 0,001$ ) que las estimaciones de ACOV ( $\bar{x} = 0,02$ ;  $dt = 0,19$ ;  $s / \sqrt{n} = 0,003$ ), con  $\eta^2 = 0,27$ . Las estimaciones de PLS-PM del parámetro  $\gamma_{11}$  presentaron un sesgo negativo en el rango del 28% al 38% en los distintos niveles de asimetría, en cambio las estimaciones ACOV presentaron un sesgo cercano a 0, dentro del rango aceptable, en los tres niveles de asimetría ( $\eta^2 = 0,18$ ). Respecto del parámetro  $\beta_{21}$ , ACOV presentó un sesgo cercano a 0, dentro del rango aceptable en los tres niveles de asimetría, en cambio PLS-PM presentó un sesgo negativo en el rango del 33% al 42% en los distintos niveles de asimetría ( $\eta^2 = 0,13$ ).

## **Discusión y Conclusiones del Estudio 1**

El presente estudio evalúa la eficiencia y el sesgo en la recuperación de parámetros de las dos técnicas de MEE más extendidas: ACOV y PLS-PM. Al analizar un modelo simple definido por tres variables latentes, cada una medida mediante 3 variables manifiestas reflectivas con escala ordinal, en distintas condiciones de especificación del modelo, tamaño muestral y nivel de asimetría habituales en psicología, educación (Micceri, 1989) y economía (Hulland, Ryan y Rayner, 2005, citados por Hwang y cols., 2010a). Utilizando además algunos valores extremos en la búsqueda de superar limitaciones de estudios previos (Hwang y cols., 2010a).

El principal aporte del presente estudio se encuentra en la utilización de variables manifiestas con escala ordinal, utilizando valores extremos de asimetría y tamaño muestral, y utilizando además el método de estimación ULS para ACOV, sobre la matriz de correlaciones policóricas.

Los resultados evidencian que un poco más de un séptimo del total de las estimaciones de los parámetros realizadas con ACOV presenta soluciones impropias, frente al 0,2% de PLS-PM. Por otra parte PLS-PM presenta estimaciones más eficientes de los parámetros del Modelo Estructural en niveles de asimetría nula, y cuando se utiliza la modalidad de modelo menos especificado.

En cuanto al sesgo en la estimación de los parámetros del Modelo Estructural, se puede señalar que ACOV presenta estimaciones con sesgos dentro del rango aceptable (inferior al 10%), en las dos modalidades de especificación del modelo y en los distintos niveles de asimetría. Sin embargo, PLS-PM presenta un sesgo negativo

que tiende a disminuir cuando el modelo está menos especificado y a aumentar cuando aumenta el nivel de asimetría.

El sesgo en las estimaciones ACOV de los parámetros del Modelo de Medida se encuentra dentro del rango aceptable en los distintos niveles de asimetría y tamaños muestrales. En cambio, PLS-PM presenta un sesgo positivo fuera del rango aceptable para valores de asimetría=0; el cual disminuye hasta situarse dentro del rango aceptable al aumentar el nivel de asimetría. Así también, PLS-PM presenta un sesgo positivo dentro del rango aceptable, en las estimaciones de parámetros del Modelo de Medida en los distintos tamaños muestrales. Estos resultados concuerdan el planteamiento de Chin (1995) y con los hallazgos de Barroso, Cepeda y Roldán (2010), y Hulland, Ryan y Rayner (2010), quienes trabajando con datos continuos y modelos más complejos, evidenciaron la tendencia de PLS-PM de subestimar los parámetros del Modelo de Estructural y a sobreestimar los parámetros del Modelo de Medida.

En el Modelo Estructural ACOV presentó una menor distancia entre sus estimaciones y el valor poblacional de los parámetros (ECM) que PLS-PM en los distintos niveles de asimetría, en las dos modalidades de especificación del modelo y en los distintos tamaños muestrales utilizados en este estudio.

En el Modelo Estructural ACOV presentó una menor diferencia absoluta promedio entre las estimaciones y los parámetros poblacionales, expresada en la misma escala de los parámetros (MAD), que PLS-PM en ambas modalidades de especificación del modelo, cuando se utilizaron los tres niveles de asimetría. Así también el promedio del MAD de las estimaciones de ACOV fue menor que el promedio del MAD de PLS-PM cuando se utilizaron tamaños muestrales muy grandes,

grandes y medianos. Cuando se utilizó un tamaño muestral pequeño ambas técnicas obtuvieron valores similares de MAD.

En el Modelo de Medida ACOV presentó una menor diferencia absoluta promedio entre las estimaciones y los parámetros poblacionales, expresada en la misma escala de los parámetros (MAD), que PLS-PM en condición de asimetría nula y tamaños muestrales grande ( $n=300$ ) y muy grande ( $n=500$ ). Al aumentar el nivel de asimetría los valores de MAD de PLS-PM se asemejan a los valores de MAD de ACOV. Con tamaños muestrales mediano y pequeño ambas técnicas presentan diferencias promedio similares en los distintos niveles de asimetría utilizados en este estudio.

La evaluación individual de la estimación de parámetros del modelo estructural, evidencia que ambas técnicas estiman el parámetro  $\gamma_{21}$  con valores superiores a 0. En torno al 10% del efecto total que recibe  $\eta_2$  en el caso de ACOV y en torno al 23,49 % en el caso de PLS-PM. Los otros dos parámetros estructurales son estimados con sesgo negativo fuera del rango aceptable con PLS-PM y con sesgo dentro del rango aceptable con ACOV.

Como conclusión se plantea que la estimación de parámetros PLS-PM es más eficiente que la estimación ACOV cuando se utiliza un modelo con variables manifiestas reflectivas con escala ordinal de cinco puntos en un modelo simple, cuando los datos presentan asimetría nula y el modelo se encuentra menos especificado. Siendo necesario considerar la presencia de un sesgo negativo (dentro del rango del 20% y 30%) para la estimación de parámetros del Modelo Estructural. Sin embargo, ACOV presenta menores diferencias entre los parámetros estimados y los parámetros poblacionales expresadas en la misma escala de los parámetros

(MAD), que PLS-PM, cuando se utilizan tamaños muestrales muy grande, grande y mediano. Con un tamaño muestral pequeño ambas técnicas presentan valores similares. Por lo que la tendencia a utilizar PLS-PM, fundamentada únicamente en la utilización de un tamaño muestral pequeño (Hair y cols. 2012; Gefen y cols. 2011, Marcoulides, 2006), no es avalada por los resultados de este estudio y no es recomendada.

Básicamente, la elección entre la utilización de ACOV o PLS-PM debe fundamentarse en la coherencia entre la finalidad del estudio y la finalidad de la técnica de análisis. Así:

1. Se recomienda utilizar PLS-PM cuando: (a). El objetivo de la investigación esté centrado en la predicción de las variables dependientes, y (b). Se requiera evitar la falta de convergencia o disminuir la probabilidad de obtención de soluciones impropias.
2. Se recomienda utilizar ACOV cuando el objetivo de la investigación esté centrado en comprobar en los datos relaciones teóricamente establecidas.

La tendencia de PLS-PM a sobreestimar el parámetro del Modelo Estructural que fue fijado a 0 en el modelo poblacional, es decir la tendencia a identificar efectos directos donde no los hay, y a presentar un sesgo negativo en los otros parámetros del Modelo Estructural, es decir la tendencia a indicar que la relación existente entre las variables latentes es menos intensa de lo que realmente es. Tendrán efectos negativos a la hora de realizar un análisis de mediación. Por lo cual, para este tipo de análisis se sugiere elegir ACOV.

Las principales limitaciones del presente estudio son la utilización de un modelo demasiado simple (con sólo tres variables latentes y tres variables manifiestas para cada variable latente, sin incluir cargas cruzadas), con valores elevados e invariantes para los parámetros, lo cual, al ser poco habitual en investigaciones aplicadas, puede restringir la generalización de los resultados encontrados. Además, se utilizan exclusivamente variables manifiestas de tipo reflectivo. Se considera que el estudio de recuperación de parámetros en estas condiciones limitadas permite establecer un punto de partida, determinando las condiciones en que ACOV y PLS-PM presentan diferencias en sus estimaciones. Así, estudios posteriores que comparen estas dos técnicas de MEE, podrán manipular otras variables experimentales, en dichas condiciones, pudiendo atribuir posibles diferencias en los resultados a las variables incorporadas.

Como sugerencia para futuros estudios sobre esta temática se sugiere incorporar alguna de las otras técnicas de MEE (p.e. GSCA), pudiendo incorporar además variables manifiestas de tipo formativo. Otros aspectos que pueden ser incluidos son: utilizar distintas escalas de medida para las variables manifiestas (p.e. ordinal y de intervalos), modelos con distintos niveles de complejidad, distintos valores de parámetros poblacionales y utilizar la modalidad de mala especificación del modelo.

## **Estudio 2: Eficiencia y sesgo en la recuperación de parámetros, con datos ordinales simulados en un modelo con variables manifiestas formativas y reflectivas, del Análisis de Estructura de Covarianza, el Modelo de Rutas mediante Mínimos Cuadrados Parciales y el Análisis Generalizado de Componentes Estructurales**

### **Introducción**

Los estudios que comparan ACOV con PLS-PM suelen utilizar datos simulados continuos (p.e. Fornell y Bookstein, 1982; Reinartz, Haenlein y Henseler, 2009). Aunque se han realizado estudios con datos en escala ordinal (p.e. Barroso, Cepeda y Roldán, 2010; Hulland, Ryan y Rayner, 2010; y Vilares, Almeida y Coehlo, 2010), éstos utilizan ACOV con el método de estimación de Máxima Verosimilitud (ML); y se limitan a la utilización de modelos con variables manifiestas reflectivas.

Considerando la diferencia entre indicadores efecto (reflectivos) e indicadores causa (formativos), en cuanto a que los primeros representan el reflejo o manifestación de un constructo (Fornell y Bookstein, 1982) y los segundos “son variables observadas que afectan directamente su variable latente” (Bollen, 2002, p. 616). En el presente estudio incorpora ambos tipos de indicadores de variables latentes en el modelo teórico utilizado. Por simplicidad, se utilizará la nomenclatura de ACOV, tanto en el diagrama de rutas como en la descripción del procedimiento, resultados y conclusiones.

Ringle, Götz, Wetzels y Wilson (2009), estudian distintos métodos de estimación de ACOV y distintos esquemas de ponderación de PLS-PM con un modelo

que incluye variables manifiestas formativas continuos en dos condiciones asimetría (0 y 2). Los resultados apuntan a que frente a un modelo con variables formativas las estimaciones de ACOV con Máxima Verosimilitud (ML) son más precisas y robustas que las estimaciones con los métodos de Mínimos Cuadrados No Ponderados (ULS), Mínimos Cuadrados Generalizados (GLS) y Asintóticamente Libre de Distribución (ADF). Por otra parte, los autores comprueban que los tres esquemas de ponderación de PLS-PM (centroide, factor y ruta) entregan las mismas estimaciones para los parámetros, lo cual concuerda con lo planteado por Tenenhaus, Esposito Vinzi, Chatelin, y Lauro, (2005). En cuanto a la comparación de ambas técnicas de MEE evidencian que las estimaciones de los parámetros formativos del modelo de medida tienden a ser sobreestimados por ACOV y subestimados por PLS-PM, los parámetros reflectivos del Modelo de Medida tienden a ser sobreestimados por PLS-PM y subestimados por ACOV. Finalmente los parámetros del Modelo Estructural tienden a ser sobrestimados por ACOV y subestimados por PLS-PM. Las limitaciones del estudio incluyen la utilización de un solo tamaño muestral ( $n=300$ ) y no se incorporan variaciones de la especificación del modelo.

Un trabajo destacable, es el realizado por Hwang, Malhotra, Kim, Tomiuk y Hong (2010a). En su estudio, realizaron simulaciones para comparar la precisión en la recuperación de parámetros ACOV, PLS-PM y GSCA con variables manifiestas continuos, en un modelo que incluye efectos cruzados. Los autores utilizaron distintos niveles de asimetría (0 y 1,25), distintos tamaños muestrales (100, 200, 300, 400 y 500) y dos niveles de especificación del modelo (bien especificado y mal especificado). Encontrando que la única condición determinante en la recuperación de parámetros fue la especificación del modelo. En concreto, cuando el modelo se encontraba correctamente especificado, ACOV recuperó los parámetros sin sesgo y recuperó mejor los parámetros que PLS-PM y GSCA. Sin embargo, cuando el modelo se encontraba incorrectamente especificado GSCA recuperó mejor los parámetros.



ACOV sobreestimó los parámetros del modelo estructural en el límite del rango aceptable (10%) y del modelo de medida fuera del rango aceptable; y PLS-PM subestimó los parámetros del modelo estructural dentro del rango aceptable y sobreestimó los parámetros del modelo medida fuera del rango aceptable.

Además, los autores describen GSCA, destacando que, junto con PLS-PM, representa una técnica basada en componentes, aunque se diferencia de PLS-PM en que cuenta con una ecuación que integra el submodelo de medida y el submodelo estructural, lo que junto con la utilización de una función de optimización global le permite proveer de una medida de ajuste global de modelo. GSCA cuenta con un único programa informático GeSCA (Hwang, 2009) utilizable de forma gratuita desde la página web en que se encuentra ([www.sem-gesca.org](http://www.sem-gesca.org)).

Entre las limitaciones de su estudio, coherentes con las sugerencias para estudios futuros, Hwang y cols. (2010a) recomiendan utilizar niveles más extremos de asimetría y tamaño muestral. Así, el presente estudio persigue profundizar en la comparación de las técnicas de MEE: ACOV, PLS-PM y GSCA, utilizando condiciones extremas de asimetría y tamaño muestral, incorporando variables manifiestas con escala ordinal, en un modelo que incluye variables manifiestas reflectivas y variables manifiestas formativas. Utilizando además dos modalidades de especificación del modelo.

## **Método**

Se diseñó un estudio de recuperación de parámetros mediante simulación de muestras sintéticas. Se consideró pertinente valorar el efecto conjunto de 4 factores: nivel de asimetría (nula, alta), tamaño muestral (pequeño, muy grande), técnica de MEE (ACOV, PLS-PM y GSCA) y especificación del modelo (más especificado y

menos especificado). Se planteó un diseño completo para la combinación de los niveles de los factores.

El diseño incluyó dos factores de comparación inter-sujetos: el efecto de la asimetría en dos niveles: 0 (nula) y 2 (alta); y el tamaño de la muestra: 50 (pequeño) y 500 (muy grande). Ambos factores inter-sujetos se combinaron de manera completa, generando muestras independientes para cada combinación de niveles.

El modelo poblacional utilizado (Figura 23) incluye tres variables latentes endógenas, tres variables manifiestas formativas que afectan a la variable latente  $\eta_1$ , la que presenta un efecto directo sobre la variable latente  $\eta_2$  y ésta presenta un efecto directo sobre la variable latente  $\eta_3$ . Este modelo es una variación del modelo utilizado por Bollen, Kirby, Curran, Paxton y Chen (2007) y Hwang y cols. (2010a) para evaluar la recuperación de parámetros con distintos métodos de estimación utilizando variables manifiestas continuas. Se excluyeron los efectos cruzados (cross-loadings) de las variables latentes hacia las variables manifiestas, para considerar una estructura lo más simple posible. Además se modificó la relación existente de las tres variables manifiestas asociadas con la primera variable latente de reflectiva a formativa.

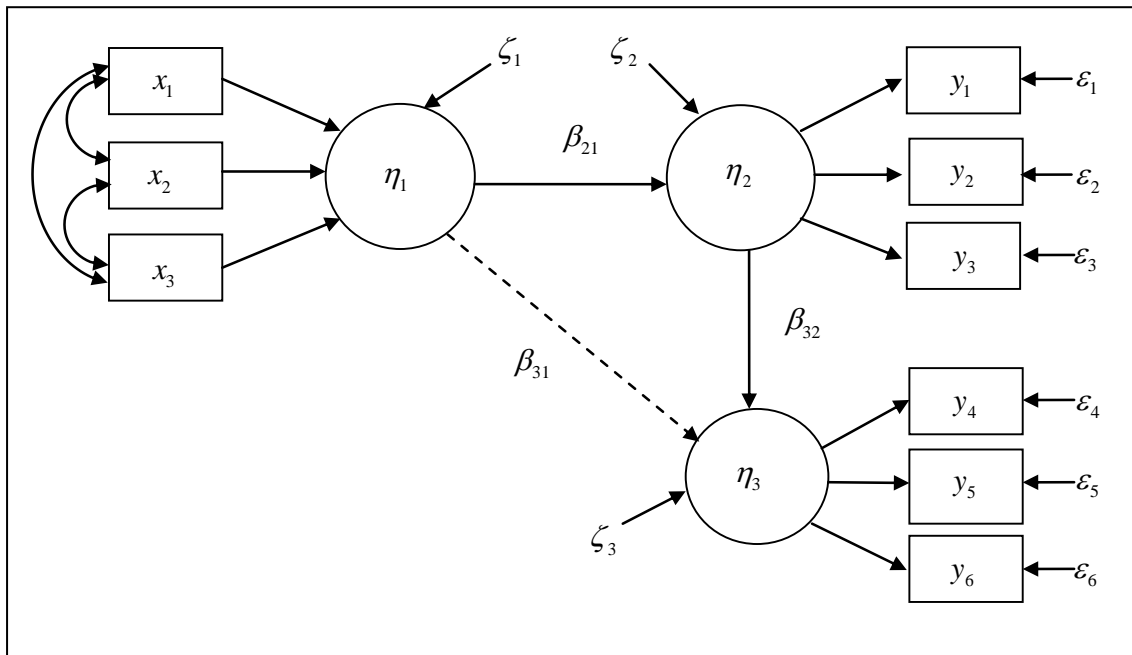


Figura 23. Diagrama de rutas del modelo poblacional, con variables manifiestas reflectivas y formativas.

\* La línea segmentada representa el parámetro que se agrega en el modelo menos especificado.

Se valoraron dos niveles de especificación del modelo en los parámetros estimables: Más especificado (se especifica el modelo poblacional) y menos especificado (se libera el parámetro  $\beta_{31}$ , el que se ha fijado a 0 para la generación de las réplicas). Este factor se definió como efecto intra-grupo.

Ambos factores intra-grupo se combinaron de manera anidada dentro de las cuatro combinaciones posibles inter-grupo. Por tanto, los 4 factores se combinan en un diseño (2x2x3x2), siendo los dos primeros factores inter-grupos y los dos segundos de medidas repetidas.

### Simulación de los datos

Se realizó una simulación Monte Carlo proactiva (Marcoulides, 2006). La que consiste en generar datos a partir de una población con valores de parámetros

hipotéticos, extrayendo muestras aleatorias repetidamente para proveer de estimaciones de los parámetros.

Se utilizó el programa Matlab 7.9.0, para generar 100 muestras para cada combinación de niveles de las condiciones experimentales inter-grupos (Anexo J), simulando un total de 400 réplicas (2 niveles de asimetría  $\times$  2 tamaños muestrales  $\times$  100 réplicas), siguiendo el procedimiento que se describe a continuación:

1.- Se definen los valores de las cargas factoriales de las variables manifiestas reflectivas ( $\Lambda_y$ ) con  $\lambda = 0,70$  y  $\lambda = 0,40$  para las variables manifiestas formativas ( $\Lambda_x$ )

2.- Se definen los valores de la matriz  $\mathbf{B} = \begin{bmatrix} 0 & 0 & 0 \\ 0,6 & 0 & 0 \\ 0 & 0,6 & 0 \end{bmatrix}$ ; y se calculan los

valores de las varianzas de los errores de predicción de las variables latentes

$\Psi = \begin{bmatrix} 0,136 & 0 & 0 \\ 0 & 0,640 & 0 \\ 0 & 0 & 0,640 \end{bmatrix}$ ; para que el valor de las varianzas de las variables

latentes sea igual a 1.

3.- Se generan las tres variables manifiestas formativas ( $x_1, x_2, x_3$ ) continuas, con distribución normal multivariante, cada una con media cero y con una correlación de 0,4 entre cada par de variables manifiestas.

4.- Se genera el error de predicción ( $\zeta_1$ ) de la variable latente  $\eta_1$ , consistente en una variable aleatoria, con media 0 y distribución normal, con la varianza calculada en el paso 2.

5.- Se calculan las varianzas de los errores de medición ( $\varepsilon$ ) de las variables manifiestas reflectivas requeridas para que las últimas tengan varianza 1.

6.- Se generan las variables latentes independientes:  $\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4, \varepsilon_5, \varepsilon_6, \zeta_2, \zeta_3$ .

7.- Se calculan las variables manifiestas reflectivas, de acuerdo con la relación de causalidad establecida en los modelos de ecuaciones estructurales (Bollen, 1989).

8. Se construye la matriz de datos  $\mathbf{X}$  con cada variable manifiesta constituyendo una columna y con tantas filas como número de sujetos se haya determinado en cada condición.

9.- Se categoriza la matriz  $\mathbf{X}$  en 5 intervalos, de acuerdo con los umbrales presentados en la Tabla 9 (extraída y modificada de Garrido, Abad y Ponsoda, 2011), para la obtención de valores ordinales tipo Likert de 5 puntos, con valores de asimetría 0 y 2 en las variables manifiestas.

Tabla 9. Umbrales utilizados para obtener variables categóricas de 5 puntos, con dos niveles de asimetría.

Asimetría	Umbrales			
	1	2	3	4
0	-1.80	-0.60	0.60	1.80
2	0.68	1.00	1.34	1.77

Cada una de las réplicas es sometida a las estimaciones pertinentes para recuperar los parámetros teóricos del modelo. De las estimaciones obtenidas, se

valoran el porcentaje de soluciones que convergen y el porcentaje de soluciones impropias, para cada una de las condiciones experimentales consideradas y para el total de muestras generadas. Se determinó un máximo de 100 iteraciones para lograr la convergencia de las soluciones y se consideró que una solución no es válida, cuando no converge o cuando la solución que converge incluye una o más varianzas negativas o incluye uno o más parámetros estimados con valor superior a 1 en valor absoluto; de acuerdo con lo planteado por Lévy, Martin y Román (2006) y con el chequeo aproximado (“*rough check*”) señalado por Bollen (1989). Las muestras que presentan soluciones no válidas son excluidas de los análisis posteriores del estudio.

Para valorar la eficiencia en la recuperación de los parámetros se utiliza la varianza promedio ( $\bar{S}^2(\hat{\theta}_{ik})$ ) de las estimaciones de los J parámetros del Modelo Estructural, del Modelo de Medida Reflectivo y del Modelo de Medida Formativo, a través de las I réplicas válidas de cada condición K. Para ello, se calcula el promedio de las diferencias al cuadrado ( $\bar{d}^2$ ) entre el parámetro estimado ( $\hat{\theta}_{ijk}$ ) y la media de las estimaciones del parámetro j ( $\bar{\hat{\theta}}_{jk}$ ) en las I réplicas válidas de cada condición intergrupo, para el Modelo Estructural, para el Modelo de Medida Reflectivo y para el Modelo de Medida Formativo:

$$\bar{d}_{jk}^2 = \frac{\sum_{i=1}^I (\hat{\theta}_{ijk} - \bar{\hat{\theta}}_{jk})^2}{I} \quad (1)$$

Luego, a través de un ANOVA de medidas repetidas se comparan las medias de las  $\bar{d}_{jk}^2$ , a través de las réplicas válidas, en cada condición experimental, correspondientes a las varianzas promedio del Modelo Estructural, a las varianzas

promedio del Modelo de Medida Reflectivo y a las varianzas promedio del Modelo de Medida Formativo:

$$\bar{S}^2(\hat{\theta}_{ik}) = \frac{\sum_{j=1}^J (\bar{d}_{jk}^2)}{J} \quad (2)$$

Valores menores de varianza promedio de los parámetros estimados implican una mayor eficiencia en la estimación de los parámetros.

Para valorar el sesgo de las estimaciones, primero se calcula el promedio de las diferencias  $(\bar{d}_{jk})$  entre el parámetro estimado y el parámetro poblacional  $(\theta_{jk})$  para el Modelo Estructural, para el Modelo de Medida Reflectivo y para el Modelo de Medida Formativo:

$$\bar{d}_{jk} = \frac{\sum_{i=1}^I (\hat{\theta}_{ijk} - \theta_{jk})}{I} \quad (3)$$

donde  $\theta_{jk}$  es el valor del parámetro poblacional  $j$  en la condición  $k$ .

Luego, a través de un ANOVA de medidas repetidas se comparan las medias del sesgo promedio de los parámetros del Modelo Estructural, del Modelo de Medida Reflectivo y del Modelo de Medida Formativo, en las réplicas válidas por condición experimental, de acuerdo con:

$$\text{Sesgo}(\theta_{ik}) = \frac{\sum_{j=1}^J (\bar{d}_{jk})}{J} \quad (4)$$

Valores de sesgo inferiores al 10% son considerados aceptables (Bollen, et al., 2007; Muthén, du Toit y Spisic, 1997; Muthén, Kaplan y Hollis, 1987). Así, se consideran aceptables valores de sesgo inferiores a 0,06 para el Modelo Estructural, valores inferiores a 0,07 para el Modelo de Medida Reflectivo y valores inferiores a 0,04 para el Modelo de Medida Formativo.

Se calculó el Error Cuadrático Medio (ECM), por ser una medida habitualmente informada en estudios de recuperación de parámetros en MEE (p.e. Asparouhou y Muthén, 2009; Boomsma y Hoogland, 2001; Hwang y cols., 2010a; Ringle, Götz, Wetzels y Wilson, 2009) y utilizada para elegir entre estimadores, ya que considera la variabilidad de los estimadores y el sesgo de los mismos en una escala equivalente (la varianza de los estimadores y el sesgo cuadrado). Para ello, primero se calcula el promedio de las diferencias al cuadrado ( $\bar{d}_j^2$ ) entre el parámetro estimado y el parámetro poblacional ( $\theta_j$ ) para el Modelo Estructural y para el Modelo de Medida, de acuerdo con:

$$\bar{d}_{jk}^2 = \frac{\sum_{i=1}^I (\hat{\theta}_{ijk} - \theta_{jk})^2}{I} \quad (5)$$

Luego, a través de un ANOVA de medidas repetidas se comparan las medias del ECM promedio del Modelo Estructural y del Modelo de Medida, en las réplicas válidas por condición experimental, de acuerdo con:

$$ECM(\hat{\theta}_{ik}) = \frac{\sum_{j=1}^J (\bar{d}_{jk}^2)}{J} \quad (6)$$



Valores menores de ECM implican una menor distancia entre los parámetros estimados y los parámetros poblacionales.

Siguiendo el planteamiento de Willmott y Matsuura (2005), quienes sugieren evitar utilizar medidas de error basadas en la suma de errores al cuadrado (p.e. RMSE), debido a la influencia que tienen sobre ellas la variabilidad de los errores. Para valorar la distancia entre los parámetros estimados y los parámetros poblacionales, se calculó la Media de las Diferencias Absolutas (MAD) entre el parámetro estimado y el parámetro poblacional, para el Modelo Estructural y del Modelo de Medida, de acuerdo con:

$$MAD(\hat{\theta}_{ik}) = \frac{\sum_{j=1}^J |\hat{\theta}_{ijk} - \theta_{jk}|}{J} \quad (7)$$

Luego, a través de un ANOVA de medidas repetidas se comparan las medias del MAD del Modelo Estructural, del Modelo de Medida Reflectivo y del Modelo de Medida Formativo, en las réplicas válidas por condición experimental. Valores menores de MAD implican una menor distancia entre la estimación del parámetro y el valor del parámetro poblacional.

## **Análisis**

Para la estimación de parámetros con ACOV se utiliza el programa LISREL v.8.8 (Jöreskog y Sörbom, 2006; Anexo K). Utilizando como método de estimación Máxima Verosimilitud (ML) sobre la matriz de correlaciones producto-momento, por ser

el método de estimación ACOV más preciso y robusto al estudiar modelos con variables formativas (Ringle, Götz, Wetzels y Wilson, 2009). Para poder estimar el modelo se reespecifican las variables manifiestas formativas como variables latentes exógenas con una sola variable manifiesta, fijando sus cargas a 1 y fijando a 0 la varianza de sus errores de medida (Williams, Edwards y Vandenberg, 2003). Además, siguiendo a MacCallum y Browne (1993), para evitar problemas de identificación, se fija a 1 el peso de una de las variables latentes exógenas, y se fija a 0 la varianza del error de estimación asociado a la variable latente afectada por las variables latentes exógenas.

Para el análisis con PLS-PM se utiliza el programa R v. 2.14.0 (R Development core team, 2011), con el paquete “plspm” (Sánchez y Trinchera, 2012; Anexo L). Considerando que estudios previos (Tenenhaus, Esposito Vinzi, Chatelin y Lauro, 2005; Ringle, Götz, Wetzels y Wilson, 2009) han comprobado la ausencia de diferencias en la estimación de parámetros al utilizar los distintos esquemas para la estimación de parámetros de PLS-PM, se opta por utilizar el esquema de rutas, ya que considera la dirección de las relaciones como han sido establecidas en el modelo de rutas predictivo (Esposito Vinzi, Trinchera y Amato, 2010).

Para la estimación de parámetros con GSCA se utiliza el único programa informático disponible para su aplicación: GeSCA (Hwang, 2009), el que se puede utilizar desde la página web: [www.sem-gesca.org](http://www.sem-gesca.org). Se analiza cada réplica desde la plataforma visual de la página web, dadas las limitaciones para la automatización de los análisis referidas por el autor del programa (Hwang, 2011).

Para la automatización de los análisis en todas las condiciones se utiliza el programa Visual FoxPro 9.0, a través del cual se ejecutan las sentencias para el análisis con LISREL (Anexo M) y R (Anexo N), y además se extraen las estimaciones

estandarizadas de los parámetros desde los archivos de resultados de los tres programas informáticos utilizados: LISREL, R y GeSCA (Anexo O, Anexo P y Anexo Q respectivamente).

Se realizan 2400 análisis de estimación (400 réplicas  $\times$  3 técnicas de MEE  $\times$  2 especificaciones del modelo). Los datos resultantes se analizan utilizando el programa IBM SPSS Statistics v.15. El modelo de análisis utilizado fue un Análisis de Varianzas (ANOVA) de medidas repetidas con 2 factores intra-grupos (técnica de MEE y especificación del modelo) y con 2 factores inter-grupos (nivel de asimetría y tamaño de la muestra), utilizando como variables dependientes la varianza de los errores de las estimaciones de los parámetros, el sesgo en las estimaciones de los parámetros, el error cuadrático medio y la media de las diferencias absolutas. Se valora el promedio por condición de los estadísticos mencionados de forma separada para los parámetros del Modelo Estructural, para los parámetros del Modelo de Medida Reflectivo y para los parámetros del Modelo de Medida Formativo.

Para evaluar la relevancia estadística de los resultados se desestimó la utilización del valor de la significación estadística observada (valor p), debido a la gran cantidad de réplicas utilizadas y la influencia que tiene sobre el valor p el tamaño de la muestra. Por dichos motivos y al igual que en los estudios de simulación de Forero, Maydeu-Olivares y Gallardo-Pujol (2009) y Hwang y cols. (2010a), se optó por considerar el tamaño del efecto como valor de comparación, teniendo en cuenta sólo los resultados con tamaño del efecto al menos mediano ( $\eta^2 \geq 0,06$ ), de acuerdo con Cohen (1988). Se informa el valor de Eta cuadrado que asume esfericidad, considerando que éste, es el mismo valor que se obtuvo para Greenhouse-Geisser, Huynh-Feldt y Límite-inferior.

## Resultados

En cuanto a la convergencia, GSCA y PLS-PM tuvieron un 100% de soluciones que convergen. Frente al 80% de soluciones que convergen de ACOV. El menor porcentaje de convergencia por condición fue de 52% para ACOV con el modelo menos especificado, asimetría de 2 y tamaño muestral de 50 (Tabla 10). Se elimina un 20% del total de réplicas por falta de convergencia.

Tabla 10. Tabla de frecuencias de réplicas que no convergen tras 100 iteraciones\*, con ACOV, PLS-PM y GSCA.

Asimetría	n	Modelo Más especificado			Modelo Menos especificado			Total		
		ACOV	PLS-PM	GSCA	ACOV	PLS-PM	GSCA	ACOV	PLS-PM	GSCA
0	50	12(88)	0(100)	0(100)	21(79)	0(100)	0(100)	25(75)	0(100)	0(100)
	500	0(100)	0(100)	0(100)	0(100)	0(100)	0(100)	54(46)	0(100)	0(100)
2	50	33(67)	0(100)	0(100)	48(52)	0(100)	0(100)	0(100)	0(100)	0(100)
	500	1(99)	0(100)	0(100)	1(99)	0(100)	0(100)	1(99)	0(100)	0(100)
Total		46(88)	0(100)	0(100)	70(82)	0(100)	0(100)	80(80)	0(100)	0(100)

\* Entre paréntesis se presenta el porcentaje de réplicas que convergen.

Respecto a la presencia de soluciones impropias, PLS-PM presentó un 3,25% de soluciones impropias, GSCA un 6% y ACOV presentó un 12,25% de soluciones impropias del total de réplicas. El mayor porcentaje de soluciones impropias por condición lo presentó ACOV con el modelo menos especificado, asimetría de 2 y un tamaño muestral de 50, con un 23%, frente al 17% de GSCA y el 11% de PLS-PM.

Las réplicas que presentaron soluciones impropias no fueron consideradas en los análisis posteriores. En total se desestimó la utilización de un 30,25% de las réplicas, considerando las 279 réplicas válidas para los análisis posteriores.

### Eficiencia

En el ANOVA de medidas repetidas (2x2x3x2) que valoraba la varianza promedio de las estimaciones estandarizadas de los parámetros del Modelo Estructural (Tabla 11.A), se evidenció un efecto de interacción de primer orden:

Tabla 11. Pruebas de efectos intra-sujetos e inter-grupo para la varianza promedio de las estimaciones estandarizadas de los parámetros del Modelo Estructural (A) y del Modelo de Medida Reflectivo (b.1) y el Modelo de Medida Formativo (b.2).

Fuente	(A) Modelo Estructural				(B) Modelo de Medida					
	GI	F	Sig	$\eta^2$	(b.1) Reflectivo		(b.2) Formativo			
					F	Sig	$\eta^2$	F	Sig	$\eta^2$
Técnica de MEE (Técnica)	2	70.7	.00	.21**	115.1	.00	.30**	49.0	.00	.15**
Especificación del modelo (Espec)	1	20.1	.00	.07**	13.5	.00	.05	3.5	.06	.01
Nivel de Asimetría (Asim)	1	24.7	.00	.08**	61.2	.00	.18**	27.7	.00	.09**
Tamaño muestral (n)	1	150.7	.00	.35**	251.2	.00	.48**	196.6	.00	.42**
Técnica*Espec	2	5.2	.00	.01	6.0	.00	.02	.9	.41	.00
Técnica*Asim	2	15.3	.00	.05	26.8	.00	.09**	19.3	.00	.07**
Técnica*n	2	58.6	.00	.18**	73.7	.00	.21**	39.3	.00	.13**
Espec*Asim	1	.2	.67	.00	14.0	.00	.05	.3	.61	.00
Espec*n	1	11.8	.00	.04	12.4	.00	.04	1.1	.29	.00
Asim * n	1	19.1	.00	.07**	35.9	.00	.12**	10.2	.00	.04
Técnica*Espec*Asim	2	.4	.68	.00	4.8	.01	.02	2.1	.12	.01
Técnica*Espec*n	2	2.9	.06	.01	6.2	.00	.02	.1	.92	.00
Técnica*Asim*n	2	11.7	.00	.04	22.1	.00	.07**	22.7	.00	.08**
Espec*Asim*n	1	.2	.69	.00	13.8	.00	.05	3.7	.06	.01
Técnica*Espec*Asim*n	2	.2	.84	.00	5.3	.01	.02	.6	.54	.00

\*\* Tamaño del efecto al menos mediano ( $\eta^2 \geq 0,06$ ).

Técnica de MEE y tamaño de la muestra ( $\eta^2=0,18$ ). Con un tamaño muestral pequeño PLS-PM presentó menores valores de varianza promedio de las estimaciones que GSCA, el cual tuvo menores valores de varianza promedio de las estimaciones que ACOV. Las tres técnicas presentaron valores pequeños y muy

similares de varianza promedio de las estimaciones cuando se utilizó un tamaño de muestra muy grande (Figura 24).

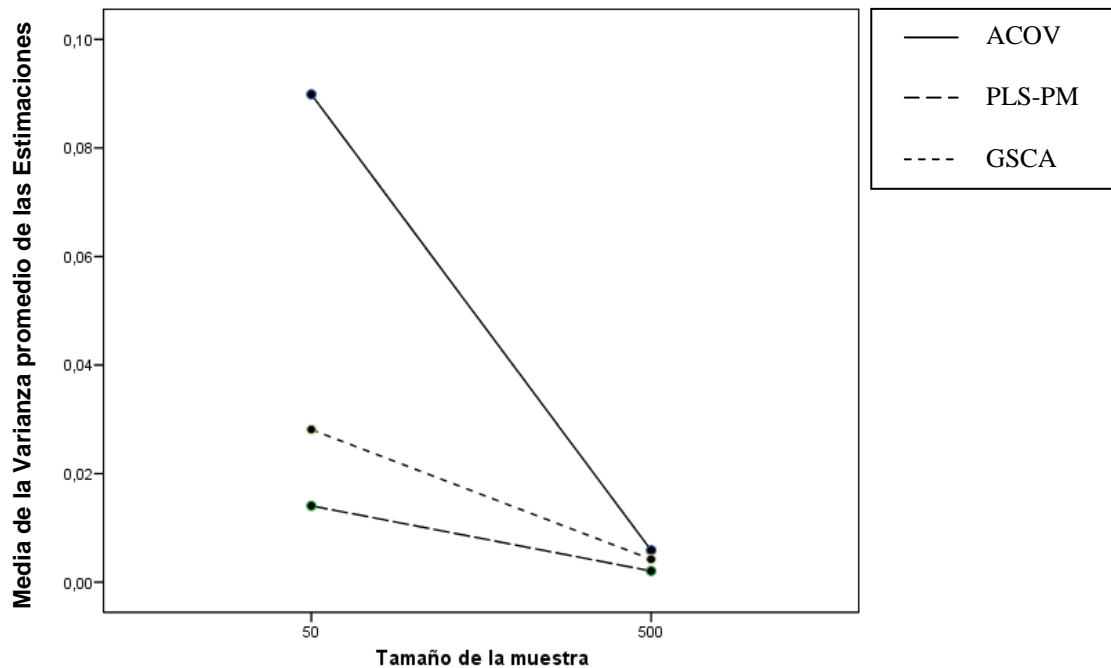


Figura 24. Medias de la varianza promedio de las estimaciones del Modelo Estructural con ACOV, PLS-PM y GSCA, según el tamaño muestral.

Nivel de asimetría y tamaño muestral ( $\eta^2=0,07$ ). Con un tamaño muestral pequeño las estimaciones con asimetría nula presentaron una menor varianza promedio que las estimaciones con asimetría alta. Las estimaciones con ambos niveles de asimetría presentaron una varianza promedio similar cuando se utilizó un tamaño de muestra muy grande.

Además se evidenció un efecto principal de la Especificación del modelo ( $\eta^2=0,07$ ). Las estimaciones con el modelo más especificado presentaron una menor varianza promedio que las estimaciones con el modelo menos especificado.

En el ANOVA de medidas repetidas (2x2x3x2) para la varianza promedio de las estimaciones de los parámetros del Modelo de Medida Reflectivo (Tabla 11.b.1), se evidenció un efecto de interacción de segundo orden entre la Técnica de MEE, tamaño de la muestra y nivel de asimetría ( $\eta^2=0,07$ ). En condiciones de asimetría nula y tamaño muestral pequeño, el valor de la varianza promedio de las estimaciones de GSCA fue menor que la varianza promedio de las estimaciones de PLS-PM y esta técnica de MEE presentó una menor varianza promedio de las estimaciones que ACOV. Con un nivel de asimetría alto, las tres técnicas de MEE aumentaron su varianza promedio de las estimaciones; GSCA continuó presentando los valores más bajos, en cambio PLS-PM aumentó a valores similares a ACOV. Las tres técnicas de MEE disminuyeron la varianza promedio de sus estimaciones a valores similares entre ellas, cuando se utilizó un tamaño muestral muy grande, tanto con asimetría nula como con asimetría alta (Figuras 25 y 26).

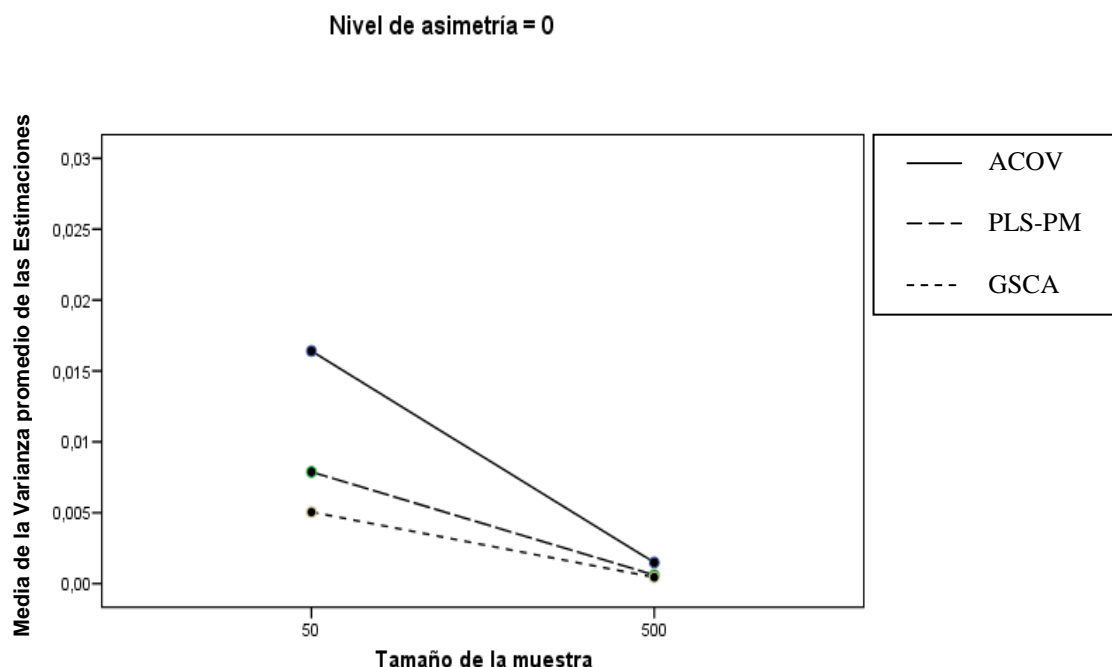


Figura 25. Medias de la varianza promedio de las estimaciones de parámetros del Modelo de Medida Reflectivo con ACOV, PLS-PM y GSCA, en nivel de asimetría nulo, según el tamaño muestral.

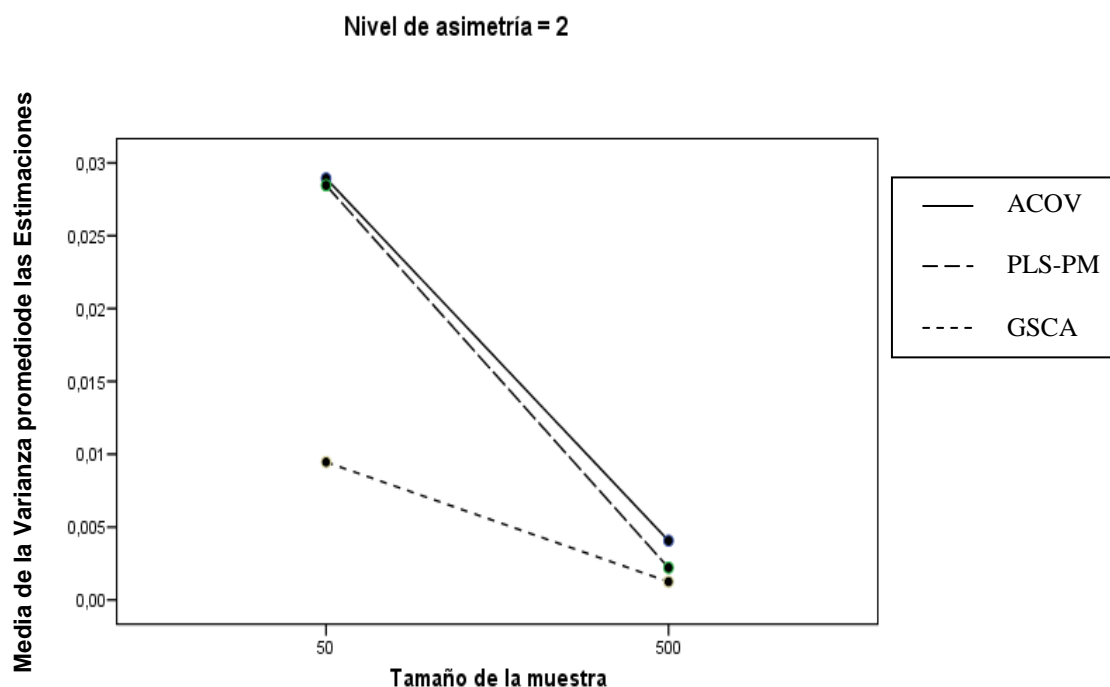


Figura 26. Medias de la varianza promedio de las estimaciones de parámetros del Modelo de Medida Reflectivo con ACOV, PLS-PM y GSCA, en nivel de asimetría alto, según el tamaño muestral.

En el ANOVA de medidas repetidas ( $2 \times 2 \times 3 \times 2$ ) para la varianza promedio de las estimaciones de los parámetros del Modelo de Medida Formativo (Tabla 11.b.2), se evidenció un efecto de interacción de segundo orden, entre la Técnica de MEE, el tamaño de la muestra y el nivel de asimetría ( $\eta^2=0,08$ ). En condiciones de asimetría nula y tamaño muestral pequeño, el valor de la varianza promedio de las estimaciones de ACOV fue menor que la varianza promedio de las estimaciones de PLS-PM y esta técnica presentó una menor varianza promedio de las estimaciones que GSCA. Con un nivel de asimetría alto ACOV y GSCA aumentaron la varianza promedio de sus estimaciones, en cambio PLS-PM se mantuvo en valores similares a los presentados cuando se utilizó asimetría nula. PLS-PM presentó una menor varianza promedio que ACOV y éste una menor varianza promedio que GSCA. Las tres técnicas de MEE disminuyeron la varianza promedio de sus estimaciones a valores similares entre ellas,



cuando se utilizó un tamaño muestral muy grande, tanto con asimetría nula como con asimetría alta (Figuras 27 y 28).

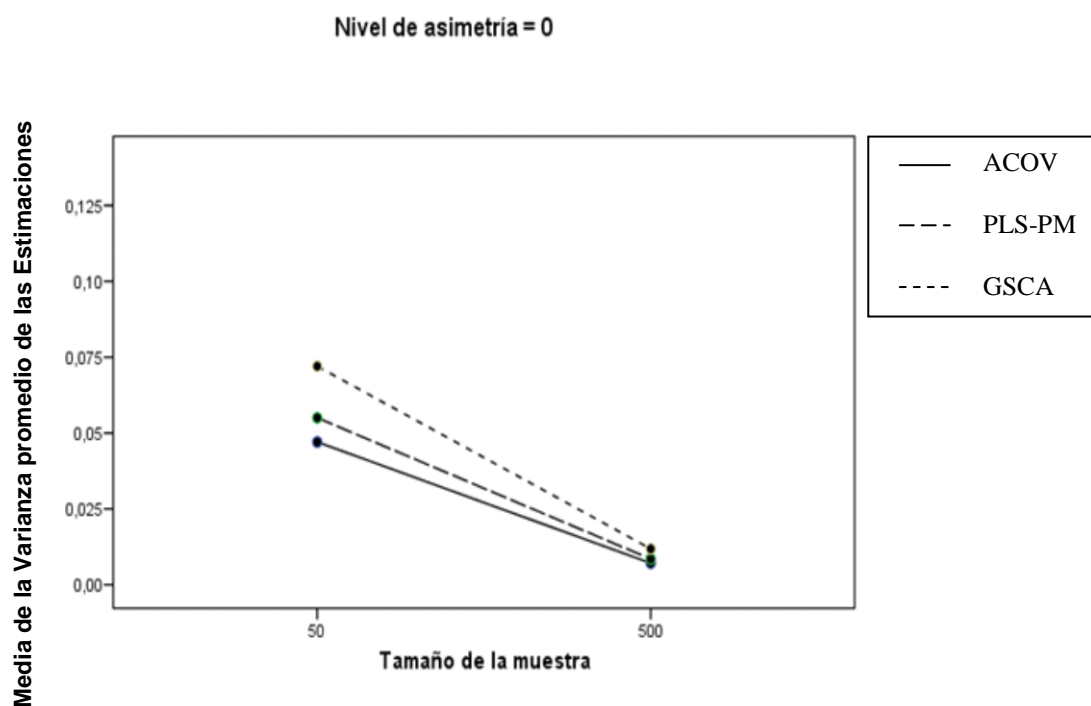


Figura 27. Medias de la varianza promedio de las estimaciones de parámetros del Modelo de Medida Formativo con ACOV, PLS-PM y GSCA, en nivel de asimetría nulo, según el tamaño muestral.

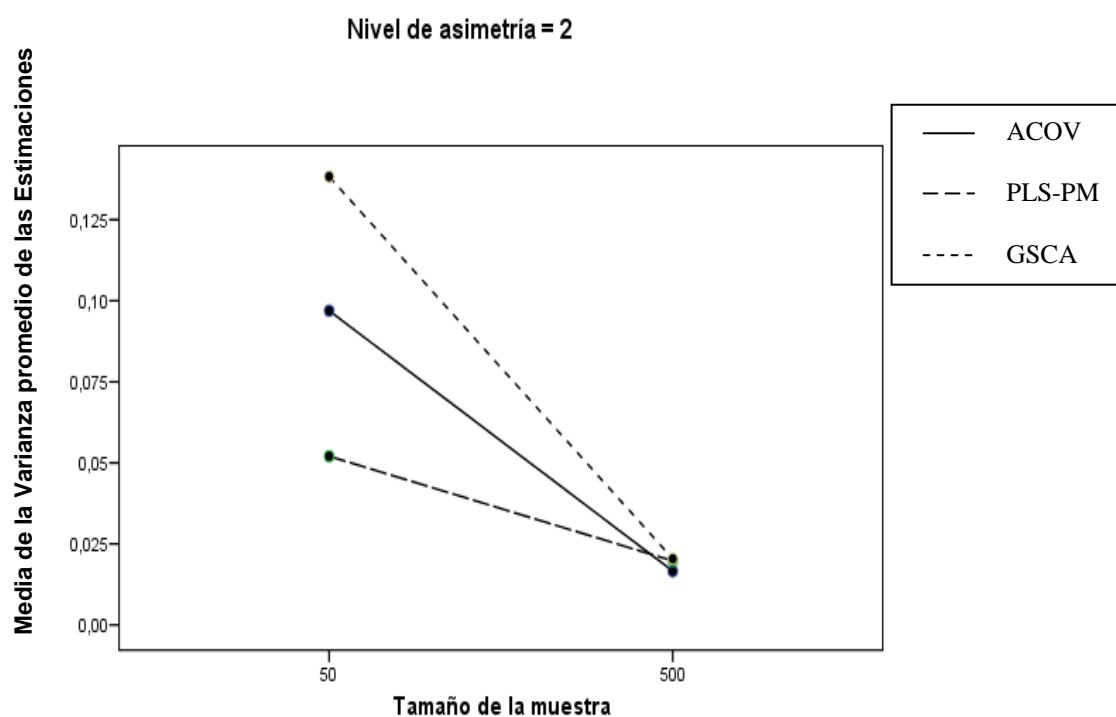


Figura 28. Medias de la varianza promedio de las estimaciones de parámetros del Modelo de Medida Formativo con ACOV, PLS-PM y GSCA, en nivel de asimetría alto, según el tamaño muestral.

Sesgo

En el ANOVA de medidas repetidas (2x2x3x4) para el sesgo promedio de las estimaciones de los parámetros del Modelo Estructural (Tabla 12.A), se evidenciaron los siguientes efectos de interacción de primer orden:

Tabla 12. Pruebas de efectos intra-sujetos e inter-grupo para el sesgo promedio de las estimaciones estandarizadas de los parámetros del modelo estructural (A) y del modelo de medida reflectivo (b.1) y formativo (b.2).

Fuente	(A) Modelo estructural				(B) Modelo de medida					
	gl	F	Sig	$\eta^2$	(b.1) Reflectivo		(b.2) Formativo			
					F	Sig	$\eta^2$	F	Sig	$\eta^2$
Técnica de MEE (Técnica)	2	472.5	.00	.63**	9303.3	.00	.97**	168.9	.00	.38**
Especificación del modelo (Espec)	1	763.9	.00	.74**	1.0	.31	.00	.5	.48	.00
Nivel de Asimetría (Asim)	1	29.5	.00	.10**	130.8	.00	.32**	4.5	.03	.02
Tamaño muestral (n)	1	32.9	.00	.11**	1.3	.25	.01	72.4	.00	.21**
Técnica*Espec	2	307.4	.00	.53**	5.0	.01	.02	2.4	.09	.01
Técnica*Asim	2	3.5	.03	.01	50.1	.00	.15**	22.9	.00	.08**
Técnica*n	2	20.3	.00	.07**	65.7	.00	.19**	44.2	.00	.14**
Espec*Asim	1	25.7	.00	.09**	14.7	.00	.05	1.8	.18	.01
Espec*n	1	4.7	.03	.02	.3	.57	.00	.00	.98	.00
Asim * n	1	8.6	.00	.03	1.7	.19	.01	6.2	.01	.02
Técnica*Espec*Asim	2	.5	.58	.00	6.4	.00	.02	.3	.75	.00
Técnica*Espec*n	2	3.3	.04	.01	11.0	.00	.04	.8	.47	.00

Técnica*Asim*n	2	9.6	.00	.03	35.0	.00	.11**	22.0	.00	.07**
Espec*Asim*n	1	1.8	.18	.01	10.8	.00	.04	.1	.77	.00
Técnica*Espec*Asi m*n	2	.5	.61	.00	5.8	.00	.02	.5	.62	.00

\*\* Tamaño del efecto al menos mediano ( $\eta^2 \geq 0,06$ ).

Técnica de MEE y especificación del modelo ( $\eta^2=0,53$ ). El sesgo promedio de las estimaciones de ACOV, fue levemente inferior a 0, encontrándose dentro del rango aceptable de sesgo, en las dos modalidades de especificación del modelo. En cambio las estimaciones de GSCA y PLS-PM presentaron un sesgo promedio negativo que se encontraba en un rango inaceptable en ambas modalidades de especificación del modelo. El sesgo promedio de las estimaciones con GSCA y PLS-PM tendió a disminuir, en torno al 15%, en condiciones de modelo menos especificado, sin alcanzar el rango aceptable (Figura 29).

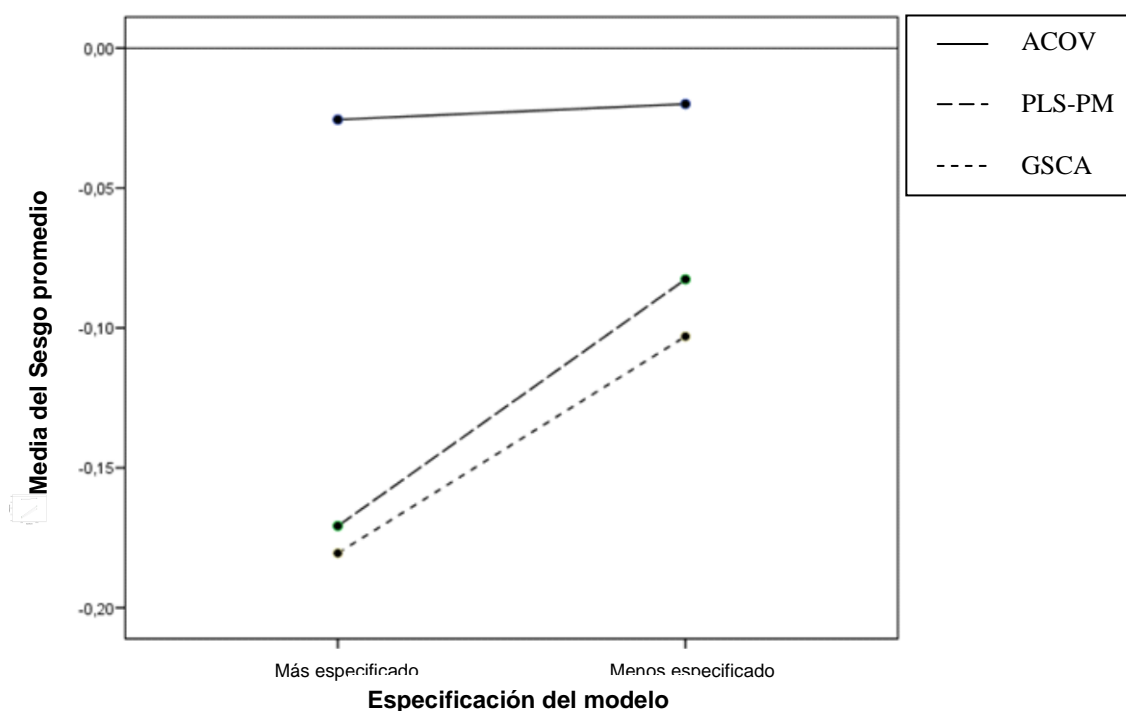


Figura 29. Medias del sesgo promedio de las estimaciones de parámetros del Modelo de Estructural con ACOV, PLS-PM y GSCA, según la especificación del modelo.

Nivel de asimetría y especificación del modelo ( $\eta^2=0,09$ ). Las estimaciones cuando se utiliza asimetría nula, con el modelo más especificado presentaron un sesgo promedio negativo, en torno al 15%, el cual es menor que el sesgo promedio de las estimaciones cuando se utilizó asimetría alta, el cual se encontró en torno al 25%, con el modelo más especificado. Cuando se utiliza el modelo menos especificado la diferencia entre los sesgos promedio con ambos niveles de asimetría disminuye. Ambos sesgos disminuyeron, llegando a niveles aceptables con asimetría nula y en torno al 13% con asimetría alta.

Técnica de MEE y tamaño de la muestra ( $\eta^2=0,07$ ). Las tres técnicas de MEE presentaron un sesgo promedio negativo, cual aumentó al aumentar el tamaño muestral. El sesgo promedio de las estimaciones de ACOV, fue levemente inferior a 0, encontrándose dentro del rango aceptable de sesgo, en las dos modalidades de tamaño muestral. En cambio las estimaciones de GSCA y PLS-PM presentaron un sesgo promedio negativo que se encontraba en un rango inaceptable en ambas modalidades de tamaño muestral; con valores similares, en torno al 25%, con tamaño muestral muy grande; y un menor sesgo promedio para PLS-PM, en torno al 15%, con tamaño muestral pequeño, frente al 20% de GSCA (Figura 30).

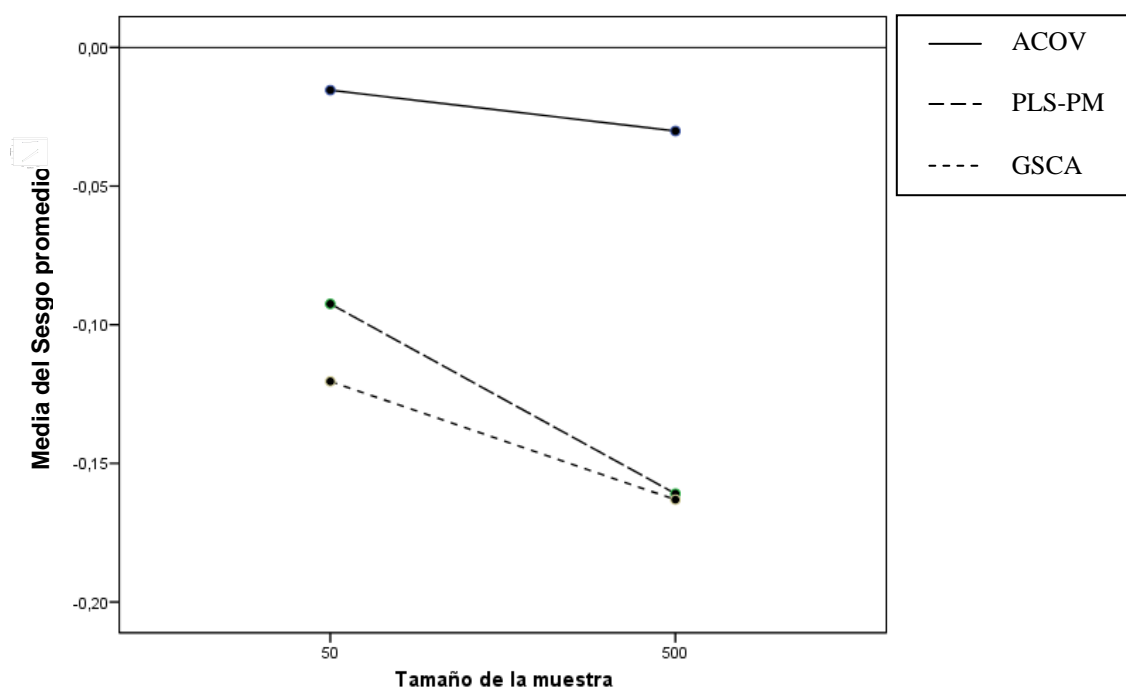


Figura 30. Medias del sesgo promedio de las estimaciones de parámetros del Modelo Estructural con ACOV, PLS-PM y GSCA, según el tamaño de la muestra.

En el ANOVA de medidas repetidas (2x2x3x2) para el sesgo promedio de las estimaciones de los parámetros del Modelo de Medida Reflectivo (Tabla 12.b.1), se evidenció un efecto de interacción de segundo orden, entre la Técnica de MEE, el tamaño de la muestra y el nivel de asimetría ( $\eta^2=0,11$ ). En condiciones de asimetría nula, PLS-PM y GSCA presentaron un sesgo promedio positivo similar (en torno al 11,50%) fuera del rango aceptable, en ambos niveles de tamaño muestral. Por su parte ACOV presentó un sesgo promedio negativo dentro del rango aceptable en ambos niveles de tamaño muestral. Cuando se utilizó una asimetría alta, PLS-PM y GSCA presentaron un sesgo promedio positivo dentro del rango aceptable en ambos niveles de tamaño muestral. En cambio ACOV presentó un sesgo promedio negativo fuera del rango aceptable en ambos niveles de tamaño muestral, el cual tendió a aumentar con un tamaño de muestra muy grande (Figuras 31 y 32).

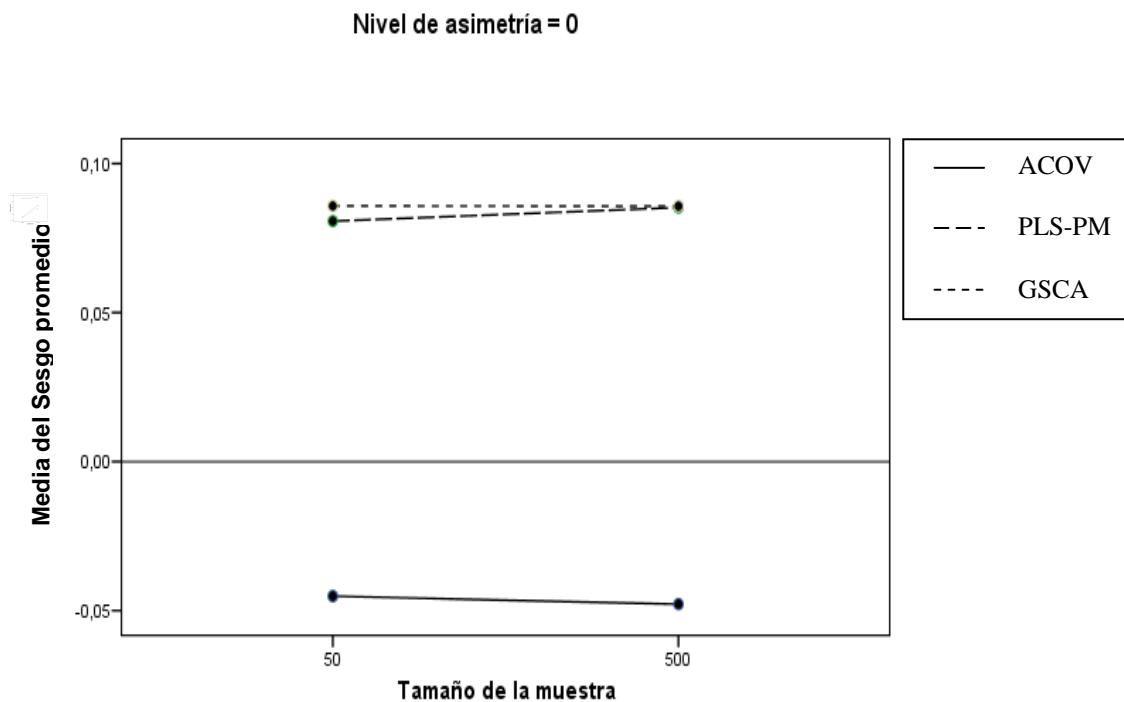


Figura 31. Medias del sesgo promedio de las estimaciones de parámetros del Modelo de Medida Reflectivo con ACOV, PLS-PM y GSCA, en nivel de asimetría nulo, según el tamaño muestral.

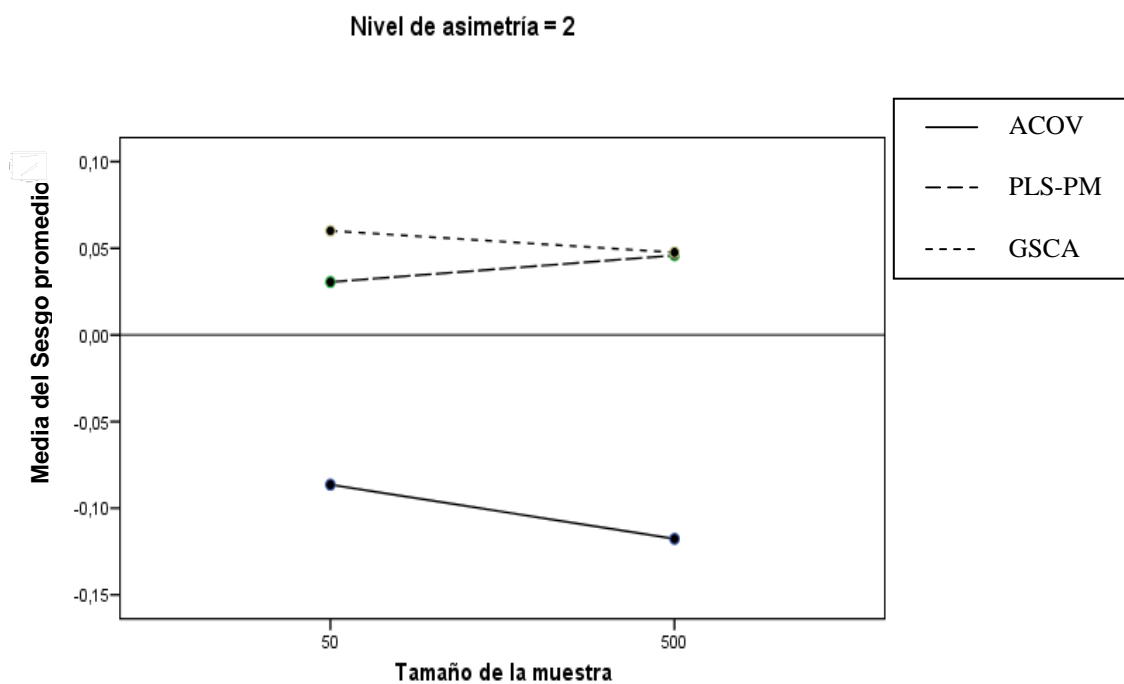


Figura 32. Medias del sesgo promedio de las estimaciones de parámetros del Modelo de Medida Reflectivo con ACOV, PLS-PM y GSCA, en nivel de asimetría alto, según el tamaño muestral.

En el ANOVA de medidas repetidas (2x2x3x2) para el sesgo promedio de las estimaciones de los parámetros del Modelo de Medida Formativo (Tabla 12.b.2), se evidenció un efecto de interacción de segundo orden entre la Técnica de MEE, el tamaño de la muestra y el nivel de asimetría ( $\eta^2=0,07$ ). En condiciones de asimetría nula, las tres técnicas de MEE presentaron un sesgo promedio que se encontró dentro del rango aceptable (10%) en ambos tamaños muestrales. El sesgo promedio positivo de PLS-PM y GSCA tendió a aumentar a niveles similares con un tamaño muestral muy grande y el sesgo promedio negativo de ACOV tendió a disminuir al aumentar el tamaño de la muestra, llegando a valores positivos cercanos a 0. En condiciones de asimetría alta las tres técnicas de MEE aumentaron su sesgo promedio en ambos tamaños muestrales, en relación con el sesgo promedio en asimetría nula. PLS-PM aumento su sesgo promedio positivo fuera del rango aceptable, en torno al 20%, con una muestra pequeña, el cual disminuyó a cerca del 14% de sesgo promedio con un tamaño muestral muy grande, condición en que GSCA presenta un sesgo promedio similar. GSCA presenta un sesgo promedio negativo dentro del rango aceptable con un tamaño muestral pequeño. ACOV presentó un sesgo promedio positivo dentro del rango aceptable cuando se utilizó un tamaño muestral muy grande; y un sesgo promedio negativo fuera del rango aceptable con una muestra pequeña (Figuras 33 y 34).

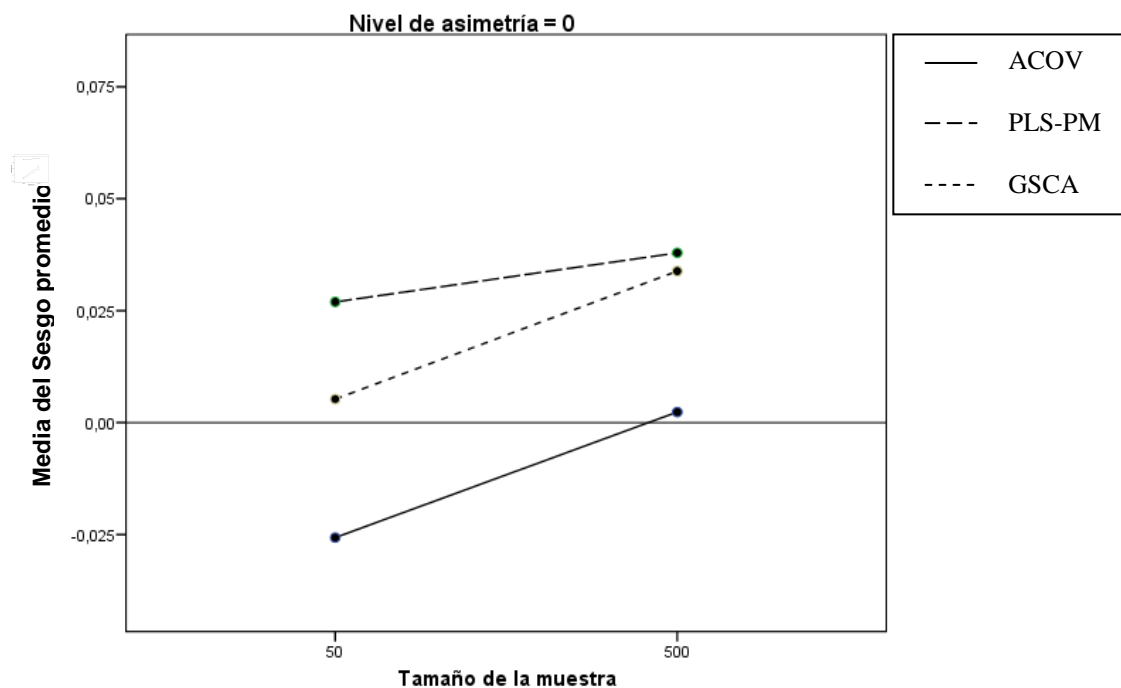


Figura 33. Medias del sesgo promedio de las estimaciones de parámetros Modelo de Medida Formativo con ACOV, PLS-PM y GSCA, en nivel de asimetría nulo, según el tamaño muestral.

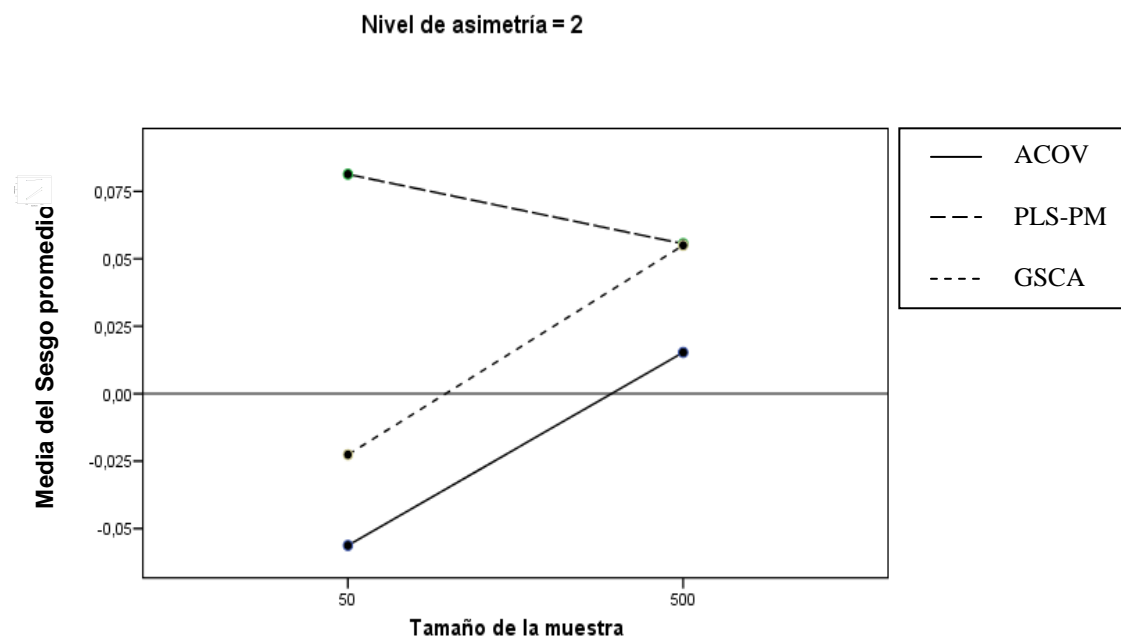


Figura 34. Medias del sesgo promedio de las estimaciones de parámetros del Modelo de Medida Formativo con ACOV, PLS-PM y GSCA, en nivel de asimetría alto, según el tamaño muestral.



*Error Cuadrático Medio (ECM)*

En el ANOVA de medidas repetidas (2x2x3x4) que valoraba el ECM promedio de las estimaciones estandarizadas de los parámetros del Modelo Estructural (Tabla 13.A), la interacción de tercer orden no presentó un tamaño del efecto mediano o superior.

Tabla 13. Pruebas de efectos intra-sujetos e inter-grupo para el ECM promedio de las estimaciones estandarizadas de los parámetros del Modelo Estructural (A) y del Modelo de Medida Reflectivo (b.1) y el Modelo de Medida Formativo (b.2).

Fuente	gl	(A) Modelo Estructural		(B) Modelo de Medida			
		F	$\eta^2$	(b.1) Reflectivo		(b.2) Formativo	
				F	$\eta^2$	F	$\eta^2$
Técnica de MEE (Técnica)	1	2.2	.01	96.0	.26**	39.4	.13**
Especificación del modelo (Espec)	1	42.2	.13**	10.9	.04	3.7	.01
Nivel de Asimetría (Asim)	2	55.5	.17**	53.1	.16**	33.3	.11**
Tamaño muestral (n)	3	77.2	.22**	207.6	.43**	198.6	.42**
Técnica*Espec	1	2.8	.01	5.8	.02	1.1	.00
Técnica*Asim	2	7.2	.03	76.4	.22**	15.9	.06**
Técnica*n	3	79.9	.23**	27.4	.09**	32.1	.10**
Espec*Asim	2	.0	.00	13.6	.05	.4	.00
Espec*n	3	55.5	.17**	9.4	.03	1.9	.01
Asim * n	6	4.7	.02	25.2	.08**	11.3	.04
Técnica*Espec*Asim	2	2.4	.01	4.8	.02	1.7	.01
Técnica*Espec*n	3	.2	.00	5.3	.02	.2	.00
Técnica*Asim*n	6	22.6	.08**	15.9	.06**	19.2	.07**
Espec*Asim*n	6	3.6	.01	11.3	.04	2.9	.01
Técnica*Espec*Asim*n	6	.3	.00	5.3	.02	.5	.00

\*\* Tamaño del efecto al menos mediano ( $\eta^2 \geq 0,06$ ).

Se evidenció un efecto de interacción de segundo orden entre la Técnica de MEE, el tamaño muestral y el nivel de asimetría ( $\eta^2=0,08$ ). Cuando se utilizó un tamaño muestral muy grande ACOV presentó un ECM promedio menor que PLS-PM y GSCA, los que presentaron valores similares. Dicha diferencia es similar en ambos niveles de asimetría. Cuando se utilizó un tamaño muestral pequeño las diferencias fueron pequeñas entre las tres técnicas de MEE en condición de asimetría nula. En condición de asimetría alta ACOV presenta un mayor ECM promedio que GSCA, técnica que presenta un valor de ECM promedio levemente superior a PLS-PM (Figura 35 y 36).

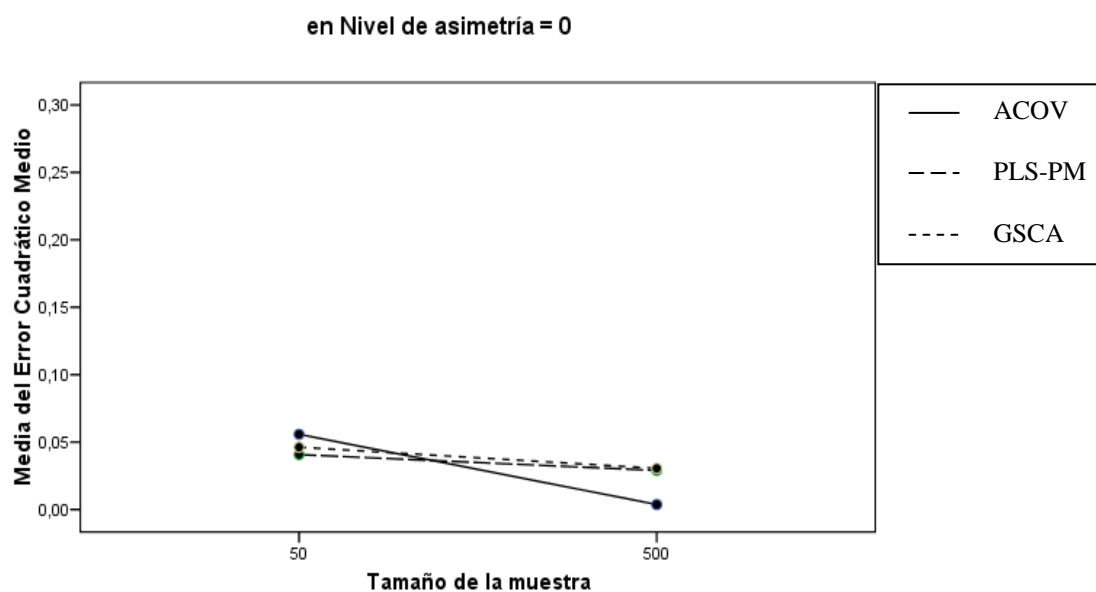


Figura 35. Medias del ECM de las estimaciones de parámetros del Modelo Estructural con ACOV, PLS-PM y GSCA, en nivel de asimetría nulo, según el tamaño muestral.

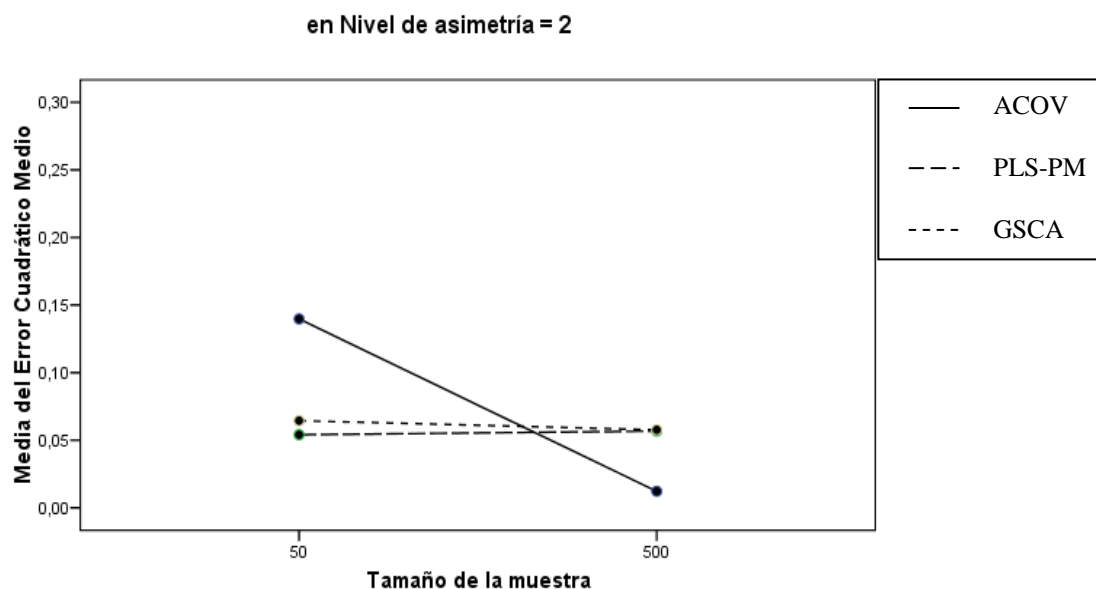


Figura 36. Medias del ECM de las estimaciones de parámetros del Modelo Estructural con ACOV, PLS-PM y GSCA, en nivel de asimetría alto, según el tamaño muestral.

Además se evidenció un efecto de interacción de primer orden entre Especificación del modelo y el tamaño muestral ( $\eta^2=0,17$ ). Cuando se utilizó un

tamaño muestral pequeño, el ECM promedio de las estimaciones con el modelo más especificado fue menor que el ECM promedio de las estimaciones con el modelo menos especificado. Cuando se utilizó un tamaño muestral muy grande el ECM promedio fue similar con ambas modalidades de especificación del modelo.

En el ANOVA de medidas repetidas (2x2x3x4) que valoraba el ECM promedio de las estimaciones estandarizadas de los parámetros del Modelo de Medida Reflectivo (Tabla 13.b.1), no se evidenció un tamaño del efecto mediano o superior para la interacción de tercer orden. Se encontró un efecto de interacción de segundo orden entre la Técnica de MEE, el tamaño muestral y el nivel de asimetría ( $\eta^2=0,06$ ). Cuando se utilizó un nivel de asimetría nulo, las tres Técnicas de MEE presentaron un ECM promedio similar con ambos tamaños muestrales. Cuando se utilizó un nivel de asimetría alto, PLS-PM y GSCA presentaron un ECM promedio similar y menor que ACOV con un tamaño muestral muy grande; con un tamaño muestral pequeño ACOV y PLS-PM presentaron un ECM promedio similar y mayor que el ECM promedio de GSCA (Figura 37 y 38).

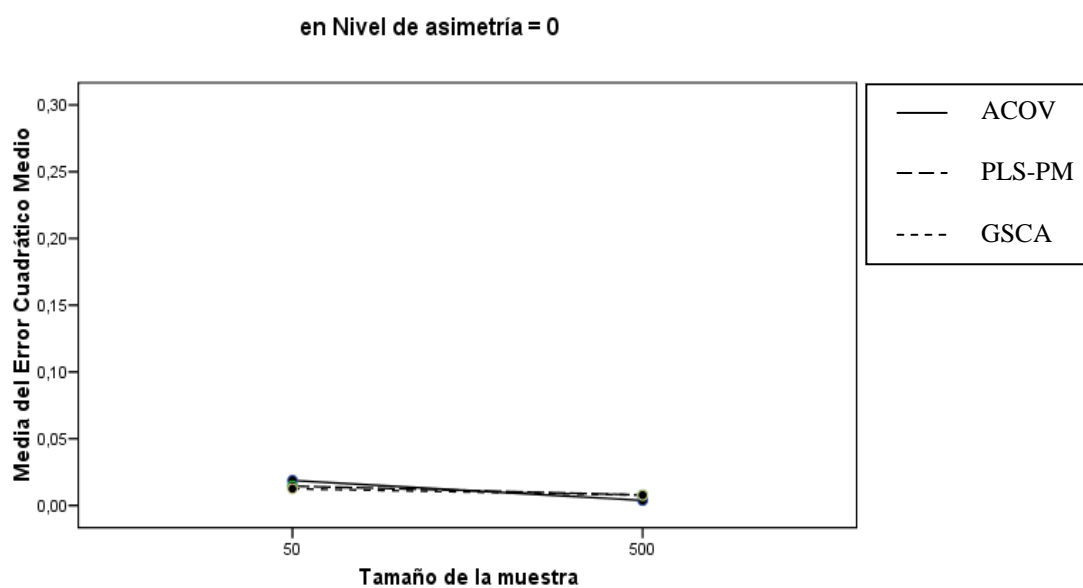


Figura 37. Medias del ECM de las estimaciones de parámetros del Modelo de Medida Reflectivo con ACOV, PLS-PM y GSCA, en nivel de asimetría nulo, según el tamaño muestral.

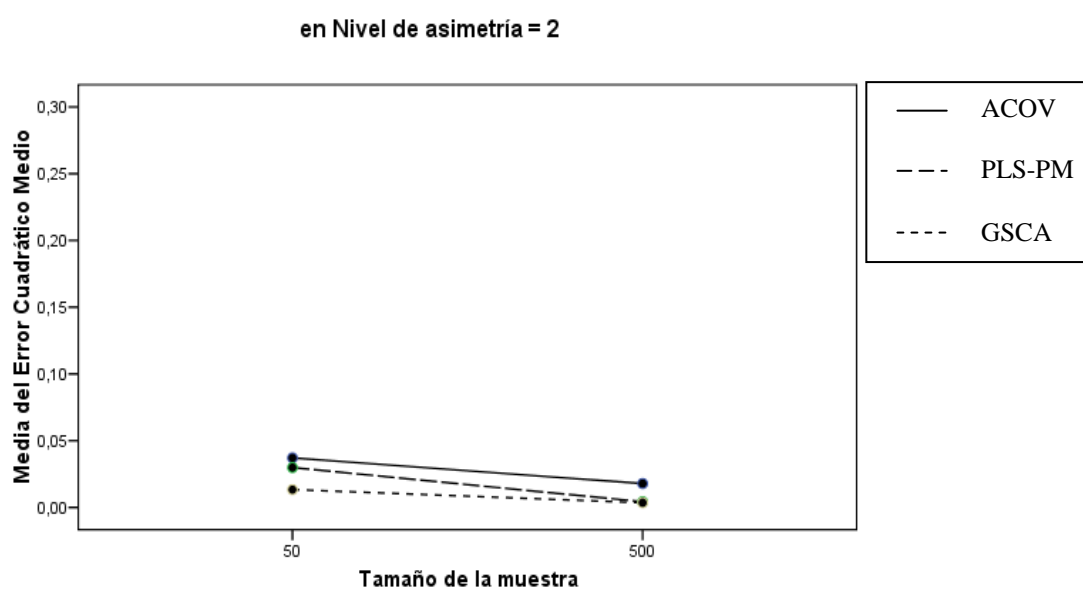


Figura 38. Medias del ECM de las estimaciones de parámetros del Modelo de Medida Reflectivo con ACOV, PLS-PM y GSCA, en nivel de asimetría alto, según el tamaño muestral.

En el ANOVA de medidas repetidas (2x2x3x4) que valoraba el ECM promedio de las estimaciones estandarizadas de los parámetros del Modelo de Medida Formativo (Tabla 13.b.2), no se evidenció un tamaño del efecto mediano o superior para la interacción de tercer orden. Se encontró un efecto de interacción de segundo orden entre la Técnica de MEE, el tamaño muestral y el nivel de asimetría ( $\eta^2=0,07$ ). Con un tamaño muestral muy grande las tres Técnicas de MEE presentaron un ECM promedio similar cuando se utilizaron ambos niveles de asimetría. Con un tamaño muestral pequeño ACOV presentó un menor ECM promedio que PLS-PM y este presentó un menor ECM promedio que GSCA en asimetría nula; al utilizar un nivel de asimetría alto PLS-PM presenta un menor ECM promedio que ACOV y GSCA, este último sigue presentando el valor más alto de ECM promedio (Figura 39 y 40).

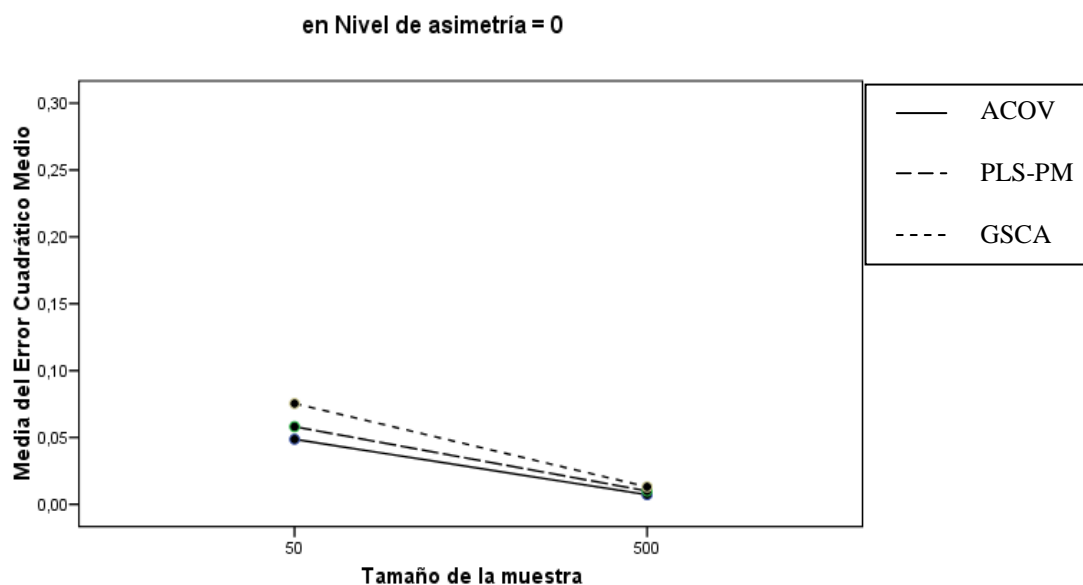


Figura 39. Medias del ECM de las estimaciones de parámetros del Modelo de Medida Formativo con ACOV, PLS-PM y GSCA, en nivel de asimetría nulo, según el tamaño muestral.

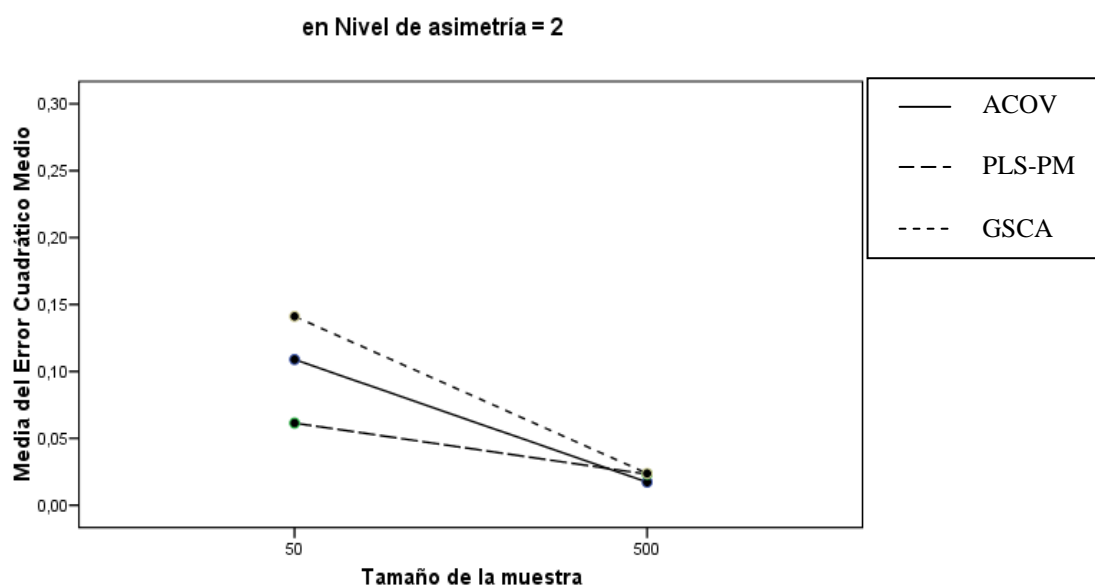


Figura 40. Medias del ECM de las estimaciones de parámetros del Modelo de Medida Formativo con ACOV, PLS-PM y GSCA, en nivel de asimetría alto, según el tamaño muestral.

#### *Media de las Diferencias Absolutas (MAD)*

En el ANOVA de medidas repetidas (2x2x3x4) que valoraba el MAD de las estimaciones estandarizadas de los parámetros del Modelo Estructural (Tabla 14.A), no se evidenció una interacción de tercer orden con tamaños del efecto medianos o superiores. Se encontró un efecto de interacción de segundo orden entre la Técnica de MEE, tamaño muestral y nivel de asimetría ( $\eta^2=0,10$ ).

Tabla 14. Pruebas de efectos intra-sujetos e inter-grupo para el promedio del MAD de las estimaciones estandarizadas de los parámetros del Modelo Estructural (A), del Modelo de Medida Reflectivo (b.1) y del Modelo de Medida Formativo (b.2).

Fuente	gl	(A) Modelo Estructural		(B) Modelo de Medida			
		F	$\eta^2$	(b.1) Reflectivo		(b.2) Formativo	
				F	$\eta^2$	F	$\eta^2$
Técnica de MEE (Técnica)	1	112.7	.29**	56.3	.17**	58.2	.18**
Especificación del modelo (Espec)	1	50.4	.16**	2.9	.01	3.1	.01
Nivel de Asimetría (Asim)	2	85.9	.24**	30.5	.10**	23.7	.08**
Tamaño muestral (n)	3	60.5	.18**	331.1	.55**	214.0	.44**
Técnica*Espec	1	19.2	.07**	3.7	.01	.4	.00
Técnica*Asim	2	9.8	.03	147.9	.35**	17.3	.06**
Técnica*n	3	294.7	.52**	9.4	.03	30.3	.10**
Espec*Asim	2	1.3	.01	2.2	.01	.5	.00
Espec*n	3	106.3	.28**	1.5	.01	2.8	.01
Asim * n	6	.1	.00	23.0	.08**	.1	.00
Técnica*Espec*Asim	2	4.4	.02	3.0	.01	5.4	.02
Técnica*Espec*n	3	4.1	.02	2.3	.01	.0	.00
Técnica*Asim*n	6	29.6	.10**	33.0	.11**	19.9	.07**
Espec*Asim*n	6	5.0	.02	2.1	.01	5.3	.02
Técnica*Espec*Asim*n	6	1.8	.01	2.1	.01	.8	.00

\*\* Tamaño del efecto al menos mediano ( $\eta^2 \geq 0,06$ ).

El promedio del MAD para las estimaciones de ACOV fue menor en ambas modalidades de asimetría cuando se utilizó un tamaño muestral muy grande. Cuando se utilizó un tamaño muestral pequeño, las tres técnicas de MEE evidenciaron valores promedio similares de MAD en condiciones de asimetría nula. En condiciones de asimetría alta PLS-PM y GSCA presentaron valores similares e inferiores a ACOV (Figura 41 y 42).



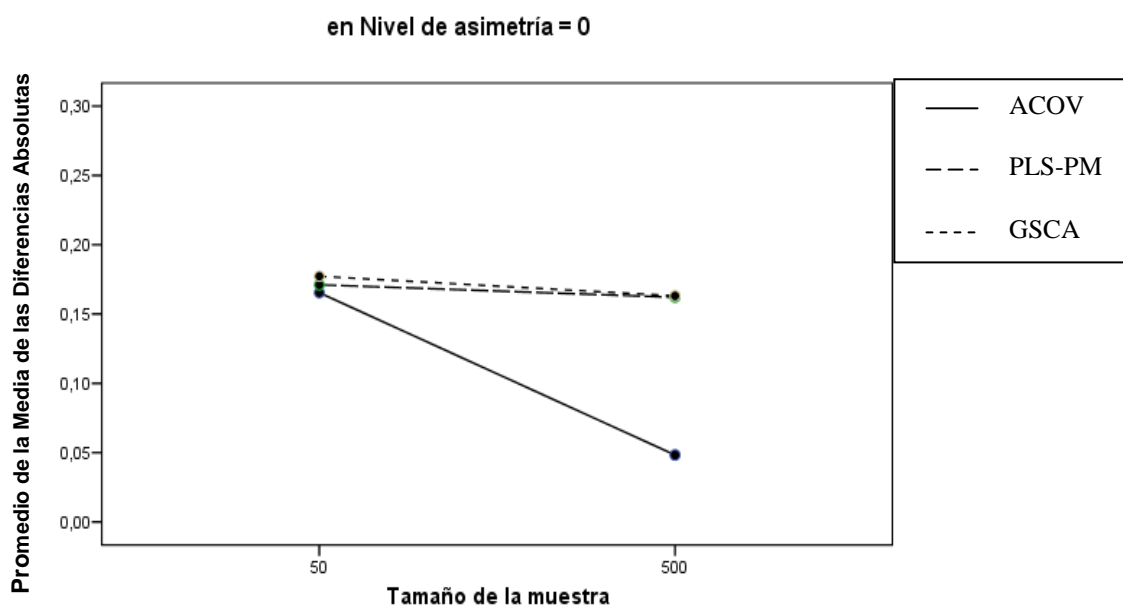


Figura 41. Promedios del MAD de las estimaciones del Modelo Estructural con ACOV, PLS-PM y GSCA, de acuerdo al tamaño de la muestra, en asimetría nula.

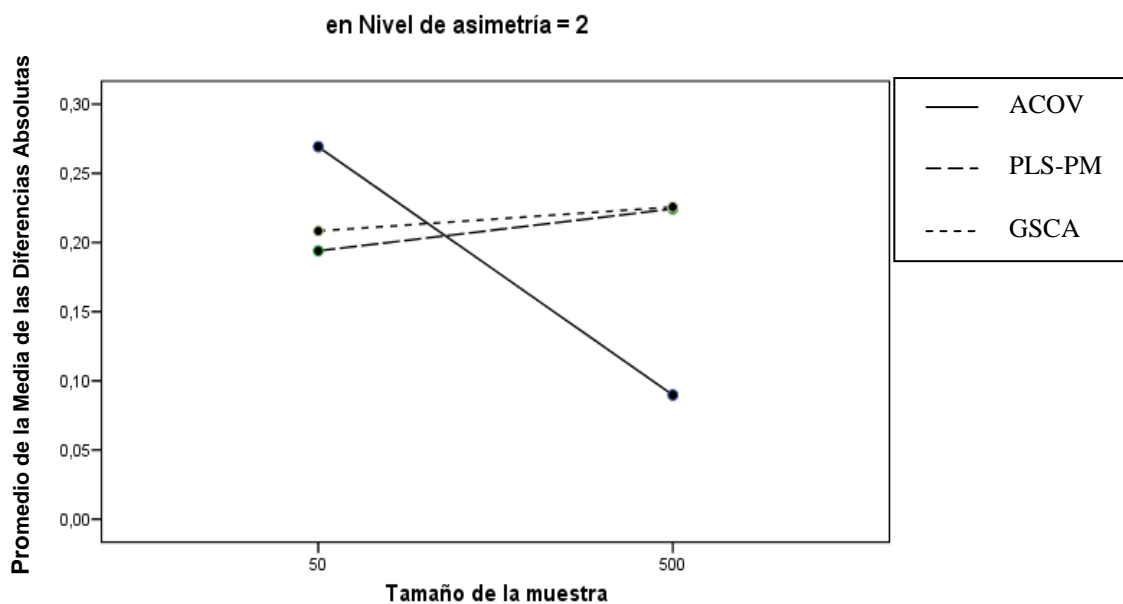


Figura 42. Promedios del MAD de las estimaciones del Modelo Estructural con ACOV, PLS-PM y GSCA, de acuerdo al tamaño de la muestra, en asimetría alta.

Además se evidenciaron los siguientes efectos de interacción de primer orden:

Técnica de MEE y especificación del modelo ( $\eta^2=0,07$ ). Las estimaciones de ACOV presentaron un menor promedio de MAD que PLS-PM y GSCA en ambas modalidades de especificación del modelo. La diferencia disminuye con el modelo menos especificado (Figura 43).

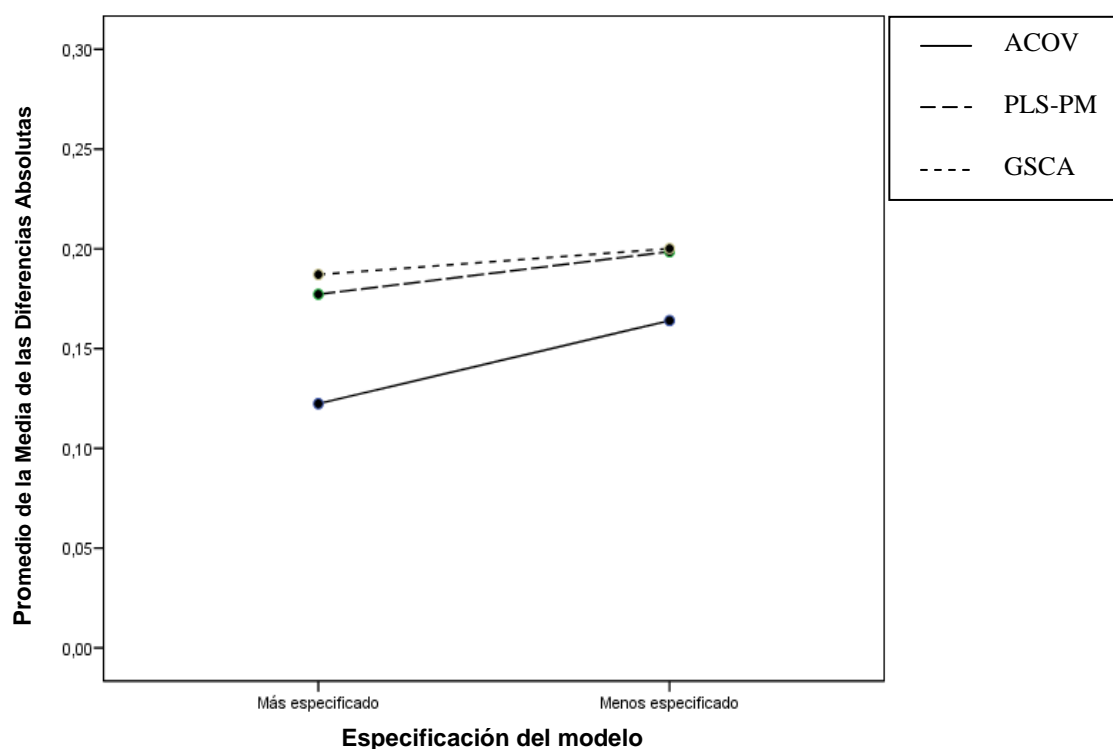


Figura 43. Promedios del MAD de las estimaciones del Modelo Estructural con ACOV, PLS-PM y GSCA, de acuerdo a la especificación del modelo.

Especificación del modelo y tamaño de la muestra ( $\eta^2=0,28$ ). El promedio del MAD de las estimaciones con el modelo más especificado es menor que el promedio del MAD con el modelo menos especificado cuando se utiliza un tamaño muestral pequeño. Con un tamaño muestral muy grande el promedio del MAD de las estimaciones con el modelo menos especificado disminuye a un valor menor que el promedio del MAD de las estimaciones con el modelo más especificado, con una

diferencia menor que la presentada entre ambas modalidades de especificación cuando se utilizó un tamaño muestral pequeño.

En el ANOVA de medidas repetidas (2x2x3x4) para el MAD de las estimaciones estandarizadas de los parámetros del Modelo de Medida Reflectivo (Tabla 14.b.1), se evidenció un efecto de interacción de segundo orden entre la Técnica de MEE, el tamaño muestral y el nivel de asimetría ( $\eta^2=0,11$ ). El promedio del MAD para las estimaciones, cuando se utilizó asimetría nula, fue menor para ACOV con un tamaño muestral muy grande, y fue similar para las tres técnicas de MEE con un tamaño muestral pequeño. Cuando se utilizó asimetría alta, con un tamaño muestral pequeño ACOV obtuvo un promedio de MAD superior que PLS-PM, y PLS-PM obtuvo un promedio de MAD superior que GSCA; con un tamaño muestral muy grande ACOV obtuvo un promedio de MAD superior que PLS-PM y GSCA, condición en que estas dos técnicas de MEE obtuvieron valores promedio de MAD similares (Figura 44 y 45).

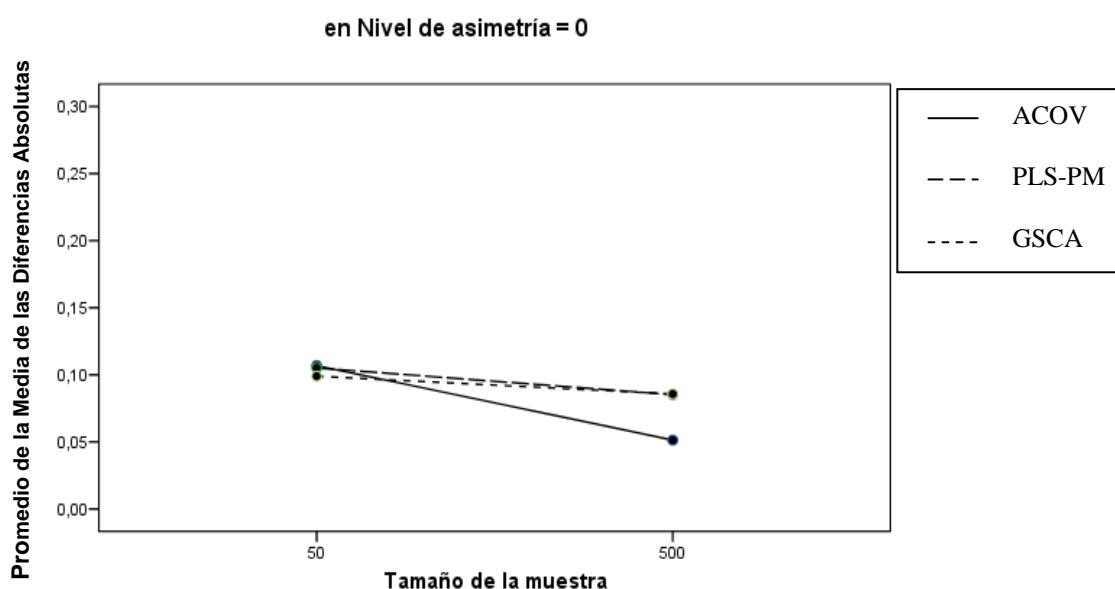


Figura 44. Promedios del MAD las estimaciones del Modelo de Medida Reflectivo con ACOV, PLS-PM y GSCA, de acuerdo al tamaño de la muestra, en asimetría nula.

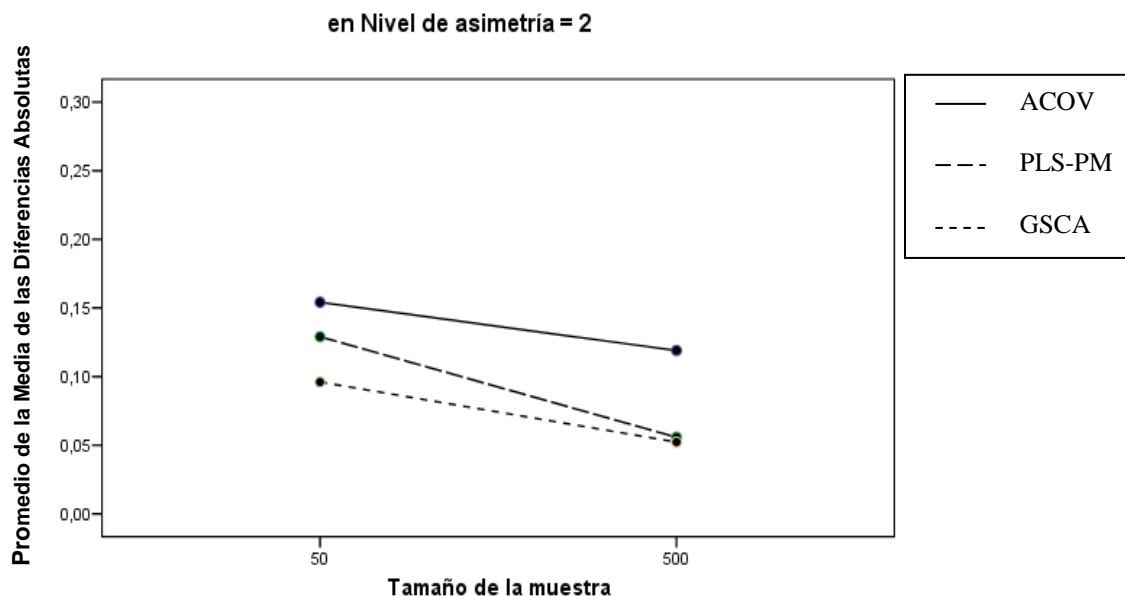


Figura 45. Promedios del MAD de las estimaciones del Modelo de Medida Reflectivo con ACOV, PLS-PM y GSCA, de acuerdo al tamaño de la muestra, en asimetría alta.

En el ANOVA de medidas repetidas (2x2x3x4) para el MAD de las estimaciones estandarizadas de los parámetros del Modelo de Medida Formativo (Tabla 14.b.2), se evidenció un efecto de interacción de segundo orden entre la Técnica de MEE, el tamaño muestral y el nivel de asimetría ( $\eta^2=0,07$ ). El promedio del MAD para las estimaciones, cuando se utilizó asimetría nula, fue menor para ACOV que para PLS-PM y GSCA con ambos tamaños muestrales. Cuando se utilizó asimetría alta, con un tamaño muestral muy grande ACOV obtuvo un promedio de MAD inferior que PLS-PM y GSCA, condición en la que ambas técnicas de MEE obtuvieron valores promedio similares de MAD; con un tamaño muestral pequeño PLS-PM obtuvo un promedio de MAD menor que ACOV y éste un promedio de MAD menor que GSCA (Figura 46 y 47).

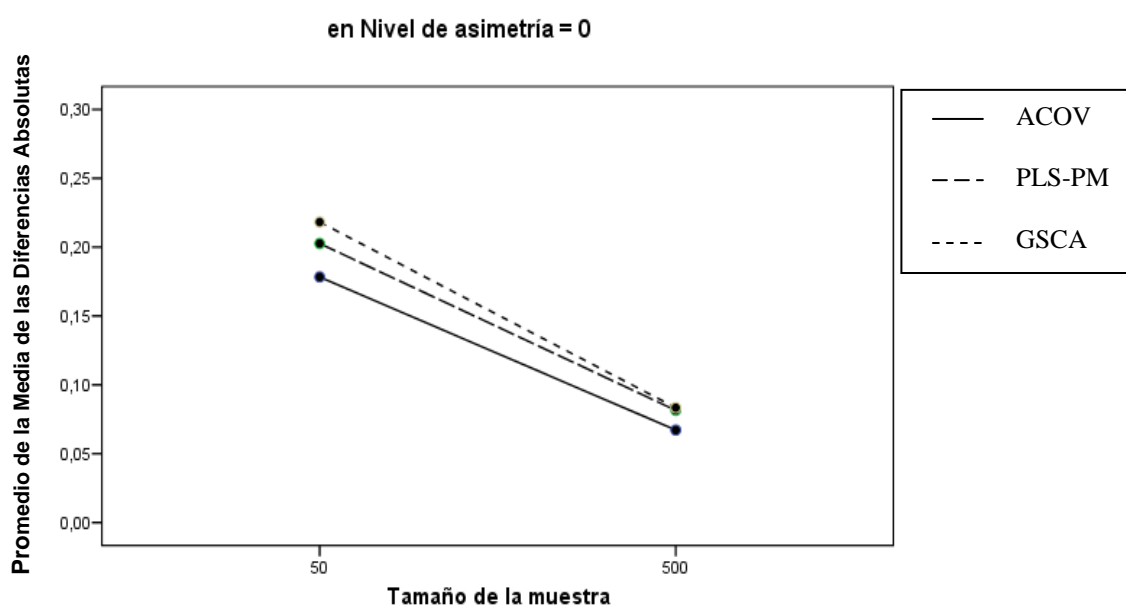


Figura 46. Promedios del MAD de las estimaciones del Modelo de Medida Formativo con ACOV, PLS-PM y GSCA, de acuerdo al tamaño de la muestra, en asimetría nula.

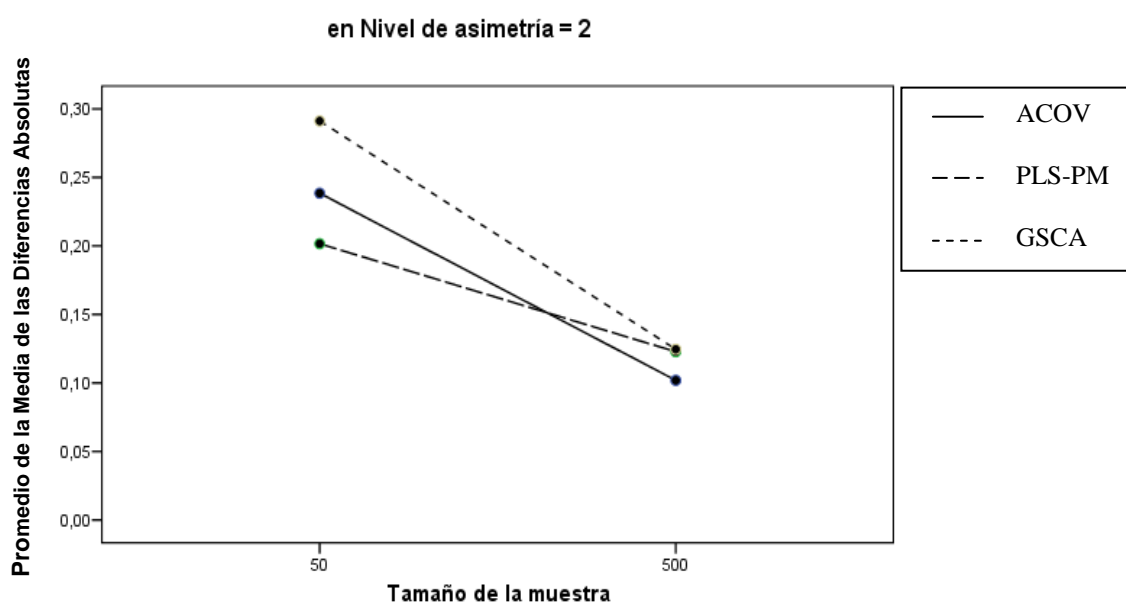


Figura 47. Promedios del MAD de las estimaciones del Modelo de Medida Formativo con ACOV, PLS-PM y GSCA, de acuerdo al tamaño de la muestra, en asimetría alta.

Por último, respecto a la incorporación del efecto impropio del parámetro estructural  $\beta_{31}$ , ACOV presentó estimaciones más pequeñas ( $\bar{x} = 0,02$ ;  $dt = 0,21$ ;  $s/\sqrt{n} = 0,013$ ) que GSCA ( $\bar{x} = 0,11$ ;  $dt = 0,12$ ;  $s/\sqrt{n} = 0,007$ ) y esta técnica de MEE presentó estimaciones más pequeñas que PLS-PM ( $\bar{x} = 0,14$ ;  $dt = 0,10$ ;  $s/\sqrt{n} = 0,006$ ), con un tamaño del efecto  $\eta^2 = 0,20$ . Por otra parte las estimaciones con una muestra muy grande presentaron una media menor ( $\bar{x} = 0,06$ ;  $dt = 0,09$ ;  $s/\sqrt{n} = 0,005$ ) que las estimaciones realizadas con una muestra pequeña ( $\bar{x} = 0,16$ ;  $dt = 0,25$ ;  $s/\sqrt{n} = 0,015$ ), con un tamaño del efecto  $\eta^2 = 0,14$ . Respecto del parámetro  $\beta_{21}$ , las estimaciones presentaron un sesgo negativo, el cual se encontró dentro del rango aceptable para ACOV; y se encontró en torno al 27% para PLS-PM y GSCA ( $\eta^2 = 0,34$ ). Respecto del parámetro  $\beta_{32}$ , las estimaciones de las tres técnicas de MEE presentaron diferencias en las medias de acuerdo al tamaño muestral y especificación del modelo ( $\eta^2 = 0,07$ ). ACOV presentó un sesgo negativo dentro del rango aceptable con el modelo más especificado, en ambos tamaños muestrales. Con el modelo menos especificado, el sesgo de ACOV siguió siendo negativo dentro del rango aceptable cuando se utilizó una muestra muy grande, pero este aumentó a niveles fuera del rango aceptable (en torno al 22%) cuando se utilizó una muestra pequeña. Por su parte PLS-PM y GSCA presentaron un sesgo negativo que se encontró fuera del rango aceptable en ambos tamaños muestrales, en ambas modalidades de especificación del modelo. Con un tamaño muestral pequeño, el sesgo de ambas técnicas de MEE aumentó de valores en torno al 28% con el modelo más especificado, a valores en torno al 45% con el modelo menos especificado. Con un tamaño muestral muy grande el sesgo aumentó de valores en torno al 37% con el modelo más especificado, a valores en torno al 42% con el modelo menos especificado.

## **Discusión y Conclusiones del Estudio 2**

El presente estudio evalúa la eficiencia y el sesgo en la recuperación de parámetros de las tres técnicas de MEE: ACOV, PLS-PM y GSCA. Al analizar un modelo que incluye variables manifiestas reflectivas y formativas medidas en escala ordinal, en dos condiciones de especificación del modelo, tamaño muestral y nivel de asimetría; habituales en psicología, educación (Micceri, 1989); y economía (Hulland, Ryan y Rayner, 2005, citados por Hwang y cols., 2010a). Utilizando además algunos valores extremos en la búsqueda de superar limitaciones de estudios previos (Hwang y cols., 2010a).

Los principales aportes del presente estudio se encuentran en la utilización de escala de medida ordinal para las variables manifiestas y la incorporación de variables manifiestas de tipo formativo, en el estudio de la recuperación de parámetros de tres técnicas de MEE, instancias poco investigadas conjuntamente en el ámbito de los MEE hasta hoy.

Los resultados evidencian que PLS-PM y GSCA convergen el 100% de las veces a diferencia del 80% de ACOV. PLS-PM presenta menos soluciones impropias que GSCA, y GSCA presenta menos soluciones impropias que ACOV. Por otra parte PLS-PM presenta estimaciones más eficientes de los parámetros del Modelo Estructural cuando se utiliza un tamaño muestral pequeño. Con tamaños muestrales muy grandes la eficiencia tiende a ser similar para las tres técnicas de MEE.

En cuanto al sesgo en la estimación de los parámetros del Modelo Estructural, se puede señalar que ACOV presenta estimaciones con sesgos dentro del rango de lo

aceptable (inferior al 10%) en las dos modalidades de especificación del modelo y en los dos tamaños muestrales utilizados. Por su parte, GSCA y PLS-PM presentan un sesgo negativo fuera del rango aceptable, en ambos tamaños muestrales y en ambos niveles de especificación del modelo.

Respecto a la eficiencia en la estimación de parámetros del Modelo de Medida Reflectivo, GSCA es más eficiente que las otras dos técnicas de MEE cuando se utiliza una muestra pequeña, en ambos niveles de asimetría. Cuando el tamaño de la muestra es muy grande, la eficiencia de las tres técnicas de MEE es similar, en ambos niveles de asimetría.

El sesgo en la estimación de parámetros del Modelo de Medida Reflectivo, es similar y positivo para PLS-PM y GSCA, en torno al 12% cuando la asimetría es nula y dentro del rango aceptable cuando la asimetría es alta. El sesgo de ACOV es negativo y se encuentra dentro del rango aceptable en los dos tamaños muestrales cuando se utiliza asimetría nula; y se encuentra en torno al 13% cuando la asimetría es alta, en ambos tamaños muestrales.

En la estimación de parámetros del Modelo de Medida Formativo, ACOV es más eficiente cuando se utiliza un tamaño muestral pequeño y asimetría nula. Las tres técnicas de MEE presentan una eficiencia similar cuando el tamaño de muestra es muy grande y la asimetría es nula. En el caso de asimetría alta, PLS-PM es más eficiente cuando se utiliza un tamaño muestral pequeño; con un tamaño muestral muy grande la eficiencia de las tres técnicas de MEE es similar.



El sesgo en la estimación de parámetros del Modelo de Medida Formativo se encuentra dentro del rango aceptable para las tres técnicas de MEE, en ambos tamaños muestrales, cuando se utiliza una asimetría nula. Con asimetría alta, PLS-PM presenta un sesgo positivo fuera del rango aceptable, en ambos tamaños muestrales. GSCA presenta un sesgo negativo dentro del rango aceptable con un tamaño muestral pequeño; y con un tamaño muestral grande presenta un sesgo positivo similar al de PLS-PM, en torno al 12%. ACOV presenta un sesgo negativo fuera del rango aceptable con un tamaño muestral pequeño el cual disminuye a valores dentro del rango aceptable con un tamaño muestral muy grande, llegando a ser positivo.

En el Modelo Estructural ACOV presentó una menor distancia entre sus estimaciones y el valor poblacional de los parámetros (ECM) que PLS-PM y GSCA cuando se utiliza un tamaño muestral muy grande, en ambos niveles de asimetría. Las tres técnicas de MEE presentan un ECM similares en condición de asimetría nula y tamaño muestral pequeño. Con asimetría alta y tamaño muestral pequeño GSCA y PLS-PM presentan valores similares e inferiores que ACOV.

En el Modelo de Medida Reflectivo las tres técnicas de MEE presentan una distancia similar entre sus estimaciones y el valor poblacional de los parámetros (ECM) cuando se utiliza asimetría nula con ambos tamaños muestrales. Con asimetría alta GSCA presenta menor ECM que ACOV y PLS-PM cuando se utiliza un tamaño muestral pequeño; Cuando se utiliza un tamaño muestral muy grande GSCA y PLS-PM presentan valores similares de ECM, los que son inferiores que los presentados por ACOV.

En el Modelo de Medida Formativo las tres técnicas de MEE presentan una distancia similar entre sus estimaciones y el valor poblacional de los parámetros (ECM) cuando se utiliza un tamaño muestral muy grande, en ambos niveles de

asimetría. Cuando se utiliza un tamaño muestral pequeño con asimetría nula ACOV presenta un menor ECM. Cuando se utiliza un tamaño muestral pequeño con asimetría alta PLS-PM presenta un menor ECM.

En el Modelo Estructural la menor diferencia promedio entre las estimaciones y los parámetros poblacionales, expresada en la misma escala de los parámetros (MAD), la presentó ACOV con un tamaño muestral muy grande, en ambos niveles de asimetría.

En el Modelo de Medida Reflectivo la menor diferencia promedio entre las estimaciones y los parámetros poblacionales, expresada en la misma escala de los parámetros (MAD), la presentó ACOV en condición de asimetría nula y tamaño muestral muy grande; en asimetría alta y tamaño muestral grande PLS-PM y GSCA presentan valores de MAD similares e inferiores que ACOV. Las tres técnicas de MEE presentan valores de MAD similares en asimetría nula y tamaño muestral pequeño. GSCA presenta menor MAD que PLS-PM y PLS-PM presenta un menor MAD que ACOV en condiciones de asimetría alta y tamaño muestral pequeño.

En el Modelo de Medida Formativo, la menor diferencia promedio entre las estimaciones y los parámetros poblacionales, expresada en la misma escala de los parámetros (MAD), la presentó ACOV cuando se utiliza asimetría nula, en ambos tamaños muestrales. En condición de asimetría alta y tamaño muestral muy grande ACOV continuó presentando un menor MAD. Sin embargo, al utilizar un tamaño muestral pequeño, con un nivel de asimetría alta, PLS-PM presentó un menor MAD que ACOV.

La evaluación individual de la estimación de los parámetros del modelo estructural, evidencia que las tres técnicas de MEE estiman el parámetro  $\beta_{31}$  con valores superiores a 0, en torno al 14,29% del efecto total que recibe  $\eta_2$  en el caso de ACOV, en torno al 26,53% en el caso de GSCA y en tono al 32,08% en el caso de PLS-PM. Los otros dos parámetros estructurales son estimados con sesgo negativo fuera del rango aceptable con GSCA y PLS-PM; y con sesgo dentro del rango aceptable con ACOV.

Como conclusión se plantea que las tres técnicas de MEE tienden a presentar una eficiencia similar cuando se utiliza un tamaño muestral muy grande, independiente del modelo (de medida o estructural) que se esté estimando. Es necesario considerar la presencia de un sesgo negativo, entre el 15% y el 25%, para las estimaciones del Modelo Estructural de PLS-PM y GSCA. Un sesgo positivo, en torno al 12%, de estas dos técnicas de MEE al estimar parámetros del Modelo de Medida Reflectivo en condición de asimetría nula. Un sesgo positivo, entre el 14% y el 20%, para PLS-PM al estimar parámetros del Modelo de Medida Formativo en condición de asimetría alta. Y un sesgo positivo, en torno al 14%, para las estimaciones de parámetros del Modelo de Medida Formativo para GSCA en condiciones de asimetría alta y tamaño muestral muy grandes.

Al utilizar un tamaño muestral pequeño, PLS-PM es más eficiente que ACOV y GSCA al estimar parámetros del Modelo Estructural. Al estimar parámetros del Modelo de Medida Reflectivo, GSCA es más eficiente. Y al estimar parámetros del Modelo de Medida Formativo, ACOV es más eficiente en asimetría nula y PLS-PM es más eficiente en asimetría alta.

ACOV presenta menores diferencias entre los parámetros estimados y los parámetros poblacionales expresadas en la misma escala de los parámetros (MAD) cuando se utiliza un tamaño muestral muy grande ( $n=500$ ) al estimar parámetros del Modelo Estructural y del Modelo de Medida Formativo. ACOV también presenta menor MAD al estimar parámetros del Modelo de Medida Reflectivo cuando se utiliza asimetría nula. Al utilizar asimetría alta PLS-PM y GSCA obtienen valores similares e inferiores de MAD, que ACOV.

PLS-PM presenta menores diferencias entre los parámetros estimados y los parámetros poblacionales expresadas en la misma escala de los parámetros (MAD) cuando se utiliza un tamaño muestral pequeño ( $n=50$ ) y asimetría alta, al estimar parámetros del Modelo Estructural y del Modelo de Medida Formativo. GSCA presenta un menor MAD al estimar parámetros del Modelo de Medida Reflectivo en las condiciones señaladas.

Las técnicas de MEE: PLS-PM y GSCA son recomendadas frente a ACOV, cuando se requiera evitar la falta de convergencia. Para disminuir la probabilidad de soluciones impropias PLS-PM es recomendada frente a GSCA y esta última es recomendada frente a ACOV. Si se busca estimar parámetros estructurales utilizando un tamaño muestral pequeño con asimetría alta, PLS-PM es la técnica más precisa para su estimación, aunque si se requiere contar con índices de bondad de ajuste del modelo a los datos GSCA es recomendado frente a ACOV.

En condiciones de tamaños muestrales muy grandes ( $n=500$ ), ACOV es la técnica recomendada, dado que presenta la menor distancia entre sus estimaciones y el valor poblacional.

Dado que GSCA cuenta con índices de ajuste global, se considera que constituye una buena alternativa frente a ACOV cuando se estiman parámetros estructurales en una muestra pequeña con asimetría alta. Se sugiere que la elección entre la utilización de ACOV, frente a PLS-PM se encuentre fundamentada en la coherencia entre la finalidad del estudio y la finalidad de la técnica de análisis. La presencia de variables manifiestas formativas no debe utilizarse como fundamento para la elección de PLS-PM o GSCA, frente a ACOV, ya que este estudio demuestra la capacidad de ACOV para obtener estimaciones más precisas que PLS-PM y GSCA, excepto cuando existe asimetría alta.

La tendencia de PLS-PM y GSCA a sobreestimar el parámetro del Modelo Estructural que fue fijado a 0 en el modelo poblacional, es decir la tendencia a identificar efectos directos donde no los hay, y a presentar un sesgo negativo en los otros parámetros del Modelo Estructural, es decir la tendencia a indicar que la relación existente entre las variables latentes es menos intensa de lo que realmente es. Tendrán efectos negativos a la hora de realizar un análisis de mediación. Por lo cual, para este tipo de análisis se sugiere elegir ACOV.

Las principales limitaciones del presente estudio son utilización de un modelo demasiado simple (con sólo tres variables latentes y tres variables manifiestas para cada variable latente, sin incluir cargas cruzadas), con valores elevados e invariantes para los parámetros, lo cual, al ser poco habitual en investigaciones aplicadas, puede restringir la generalización de los resultados encontrados. Además, la necesidad de pago para acceder a algunos de los programas informáticos utilizados, puede restringir la posibilidad de réplica del estudio. Se considera que el estudio de la recuperación de parámetros en estas condiciones limitadas permite establecer un punto de partida, determinando las condiciones en que las tres técnicas de MEE (ACOV, PLS-PM y

GSCA) presentan diferencias en las estimaciones. Así, estudios posteriores que comparen estas tres técnicas podrán manipular otras variables experimentales, en dichas condiciones, pudiendo atribuir posibles diferencias en los resultados a las variables incorporadas.

Como sugerencia para futuros estudios sobre esta temática se sugiere incorporar alguna de las otras técnicas de MEE (p.e. el método de Croon), pudiendo incorporar, modelos con distintos niveles de complejidad, distintos valores de los parámetros poblacionales, la modalidad de mala especificación del modelo, cargas cruzadas y utilizar programas informáticos gratuitos (p.e. R y GeSCA), que faciliten la réplica de los mismos.

## DISCUSIÓN GENERAL

En un modelo que presenta variables manifiestas exclusivamente reflectivas en escala ordinal de 5 puntos, las estimaciones de los parámetros del Modelo Estructural de PLS-PM son más eficientes que las estimaciones de ACOV. En un modelo con variables manifiestas reflectivas y formativas PLS-PM y GSCA presentan una eficiencia similar y mayor que ACOV, al utilizar un tamaño muestral pequeño. La diferencia en la eficiencia de PLS-PM con GSCA es menor que la diferencia en la eficiencia con ACOV, lo que sustenta la Hipótesis 1 para el Modelo Estructural. A pesar del sesgo negativo que presentan dichas estimaciones, PLS-PM y GSCA presentan estimaciones de los parámetros del Modelo Estructural más precisas que ACOV, cuando se utiliza un tamaño muestral pequeño y asimetría alta. Con tamaños muestrales muy grandes ( $n=500$ ), la eficiencia tiende a ser similar para las tres técnicas de MEE. Dichos resultados son coherentes con el requerimiento de distribución normal multivariante de ACOV, particularmente cuando se utiliza el método de estimación de ML (Bollen, 1989; Fornell y Bookstein, 1982; Jöreskog y Wold, 1982) y la ausencia de requisitos distribucionales para PLS-PM. Por lo que no se sustenta la Hipótesis 2 para el Modelo Estructural.

En cuanto a las estimaciones de los parámetros del Modelo de Medida Reflectivo, GSCA presenta estimaciones más eficientes que las estimaciones de PLS-PM y ACOV, cuando se utiliza un tamaño muestral pequeño en asimetría alta. La diferencia en la eficiencia de GSCA con PLS-PM es menor que la diferencia en la eficiencia con ACOV, en asimetría nula, lo que sustenta la Hipótesis 1 para el Modelo de Medida Reflectivo en dichas condiciones. Con asimetría alta y tamaño muestral pequeño, PLS-PM y ACOV presentan una eficiencia similar, lo que no sustenta la Hipótesis 1 para el Modelo de Medida Reflectivo en asimetría alta y tamaño muestral pequeño. Con tamaños muestrales muy grandes ( $n=500$ ), la eficiencia tiende a ser

similar para las tres técnicas de MEE, independientemente del nivel de asimetría, con lo cual no se sustenta la Hipótesis 2 para el Modelo de Medida Reflectivo.

Respecto a las estimaciones de los parámetros del Modelo de Medida Formativo, las estimaciones de ACOV son más eficientes que las estimaciones de PLS-PM y GSCA, cuando se utiliza un tamaño muestral pequeño y asimetría nula. Con lo que no se sustenta la Hipótesis 1 para el Modelo de Medida Formativo en las condiciones señaladas. Con asimetría alta y tamaño muestral pequeño, PLS-PM es más eficiente que GSCA y ACOV, y esta última técnica es más eficiente que GSCA, con lo que no se sustenta la Hipótesis 1 para el Modelo de Medida Formativo. Con tamaños muestrales muy grandes ( $n=500$ ), la eficiencia tiende a ser similar para las tres técnicas de MEE, independientemente del nivel de asimetría, con lo cual no se sustenta la Hipótesis 2 para el Modelo de Medida Reflectivo.

En cuanto a la precisión en la estimación de parámetros del Modelo Estructural, ACOV es más preciso que PLS-PM y GSCA cuando se utiliza un tamaño muestral muy grande, en ambos niveles de asimetría. ACOV, también es más preciso al estimar parámetros del Modelo de Medida Reflectivo ACOV, en condición de asimetría nula y tamaño muestral muy grande. Con lo que se sustenta la Hipótesis 4 para el Modelo Estructural.

Respecto a la precisión en la recuperación de parámetros del Modelo de Medida Reflectivo, PLS-PM y GSCA son más precisos que ACOV en condición de asimetría alta y tamaño muestral grande. Así también, en condiciones de asimetría alta GSCA es más preciso que PLS-PM y PLS-PM es más preciso que ACOV. Ambos resultados aportan sustento a la Hipótesis 3 para el Modelo de Medida Reflectivo.



En el Modelo de Medida Formativo, las tres técnicas de MEE tienden a ser más precisas con un tamaño muestral grande que con un tamaño muestral pequeño. Al utilizar un tamaño muestral pequeño las diferencias entre las tres técnicas aumenta en la medida que aumenta el nivel de asimetría. Al utilizar un tamaño muestral pequeño, con un nivel de asimetría alta, PLS-PM es más preciso que ACOV, lo que en parte sustenta la Hipótesis 3 para el Modelo de Medida Formativo, en cuanto a la mayor precisión por parte de una técnica de la aproximación basada en componentes frente a ACOV. Sin embargo, este último es más preciso que GSCA en la condición señalada.

Los niveles de especificación del modelo generan diferencias en la precisión de la estimación de parámetros del Modelo Estructural, siendo ACOV, más eficiente que PLS-PM y GSCA en ambas modalidades de especificación del modelo. Por lo que no se sustenta la Hipótesis 5. Cabe señalar que esta situación podría cambiar al utilizar la modalidad de mala especificación del modelo, acorde a lo realizado por Hwang y cols. (2010a).

En cuanto al sesgo, ACOV presenta estimaciones con sesgos dentro del rango aceptable (inferior al 10%) de los parámetros del Modelo Estructural, cuando se estiman parámetros tanto del modelo con variables manifiestas exclusivamente reflectivas (Estudio 1), como cuando el modelo incluye variables manifiestas reflectivas y formativas (Estudio 2). Así también, ACOV presenta estimaciones con sesgo dentro del rango aceptable cuando se estiman los parámetros del Modelo de Medida Reflectivo, en un modelo con variables manifiestas exclusivamente reflectivas. Sin embargo, al utilizar un modelo con variables manifiestas formativas y reflectivas, con asimetría alta, el sesgo negativo de las estimaciones de ACOV se encuentra fuera del rango aceptable.

En cuanto a la convergencia de las soluciones, ACOV presenta un elevado porcentaje de soluciones que no convergen, en comparación con las otras dos técnicas de MEE. En condiciones de tamaño muestral pequeño, asimetría alta y modelo menos especificado, cuando se estiman parámetros de un modelo con variables exclusivamente reflectivas el porcentaje de soluciones impropias de ACOV es del 17,4%. Cuando se estiman parámetros de un modelo con variables manifiestas reflectivas y formativas, es del 48%.

PLS-PM presenta un sesgo positivo, fuera del rango aceptable de 10%, en la estimación de parámetros del Modelo de Medida Reflectivo cuando se utiliza asimetría nula. El que disminuye dentro del rango aceptable al aumentar el nivel de asimetría, tanto si el modelo incluye variables manifiestas exclusivamente reflectivas o incluye variables manifiestas reflectivas y formativas.

El sesgo negativo en las estimaciones de los parámetros del Modelo Estructural de PLS-PM tiende a disminuir cuando el modelo se encuentra menos especificado, independientemente del tipo de variables manifiestas que incluya el modelo (exclusivamente reflectivas o reflectivas y formativas).

ACOV presenta estimaciones con un promedio menor del efecto impropio del parámetro fijado a 0 en la simulación de los datos, que PLS-PM y GSCA. Obteniendo valores similares al estimar el parámetro con el modelo que incluye variables manifiestas exclusivamente reflectivas, como con el modelo que incluye variables manifiestas reflectivas y formativas, a pesar de la diferencia existente en la cantidad de réplicas utilizadas en cada estudio.

La presente tesis doctoral, al utilizar simulación de datos, incluye condiciones experimentales limitadas y no exhaustivas de las condiciones posibles de ser

encontradas en contextos aplicados, lo cual restringe la generalización de los resultados sólo a contextos similares a los utilizados en esta tesis.

Además, la búsqueda de un contexto lo más simple posible que permita apreciar con claridad los efectos que generan cambios en la estimación de parámetros, conllevó la exclusión de condiciones que sería interesante investigar en estudios futuros, tales como la utilización de un modelo más complejo (con mayor cantidad de variables latentes, mayor cantidad de variables manifiestas por variable latente) que incluya cargas cruzadas y distintos valores para las cargas factoriales. Para estudio posteriores se recomienda incorporara algunas de las condiciones señaladas. Pudiendo ser interesante además incorporar el estudio de los índices de bondad de ajuste de ACOV y GSCA, e incluir la modalidad de mala especificación del modelo.



## CONCLUSIONES

La presente tesis doctoral constituye un aporte al estudio de los modelos de ecuaciones estructurales al profundizar en el estudio de la recuperación de parámetros. Partiendo del reconocimiento de distintas técnicas de MEE, evalúa la eficiencia y sesgo en la estimación de parámetros de ACOV, PLS-PM y GSCA, en condiciones que no se han utilizado conjuntamente en estudios previos y que se consideran habituales en ciencias sociales, tales como la utilización de variables manifiestas con escala ordinal, en un modelo con variables manifiestas reflectivas y formativas. Además se incorporan condiciones habitualmente utilizadas en el estudio de la recuperación de parámetros tales como distintos tamaños muestrales, niveles de asimetría y especificaciones del modelo; con niveles habituales en psicología, educación (Micceri, 1989); y economía (Hulland, Ryan y Rayner, 2005, citados por Hwang y cols., 2010a). Utilizando además algunos valores extremos ( $n=50$  y  $asimetría=2$ ) para superar limitaciones de estudios previos (Hwang y cols., 2010a).

Los resultados de los estudios de la parte empírica, en conjunto con la revisión teórica y de investigaciones previas sobre las tres técnicas de MEE consideradas en la presente tesis doctoral permiten plantear lo siguiente:

- 1.- Se recomienda utilizar ACOV cuando: (a). El objetivo del investigador este centrado en la estimación de parámetros estructurales y la comprobación del ajuste del modelo a los datos, (b). Se utilice un modelo con variables manifiestas reflectivas con escala ordinal de cinco puntos en muestras muy grandes, con asimetría nula, (c). Se utilice un modelo con variables manifiestas reflectivas y formativas, con escala ordinal de cinco puntos en

muestras muy grandes, y (d) cuando se requiera realizar un análisis de mediación entre variables estructurales.

2.- Se recomienda utilizar PLS-PM cuando: (a). El objetivo del investigador este centrado en la predicción de las variables dependientes del modelo, (b). Se utilicen variables manifiestas formativas con escala ordinal de cinco puntos en muestras pequeñas con asimetría alta, y (c). ACOV presente ausencia de convergencia o soluciones impropias, en las condiciones de especificación utilizadas en este trabajo.

Es necesario considerar la presencia de sesgo negativo de las estimaciones de PLS-PM, en el rango del 20% al 30%, para la estimación de parámetros del modelo estructural y un sesgo positivo, en torno al 10% para el Modelo de Medida.

3.- Se recomienda utilizar GSCA cuando: (a). El objetivo del investigador este centrado en la predicción de las variables dependientes y se utilice un modelo con variables manifiestas reflectivas (dependientes) y formativas (independientes) con asimetría alta en una muestra pequeña, (b). Se requiera contar con la comprobación del ajuste del modelo a los datos, y (c). ACOV o PLS-PM presenten ausencia de convergencia o soluciones impropias, en las condiciones de especificación utilizadas en este trabajo.

ACOV es la técnica recomendada frente a las otras dos técnicas de MEE. Concordando con Hwang y cols. (2010a) se considera que GSCA es una buena alternativa para la estimación de parámetros cuando ACOV presente ausencia de convergencia o soluciones impropias. Dada la inclusión de índices de bondad de ajuste global del modelo por parte de GSCA, que permiten evaluar el ajuste de los

modelos a los datos, lo que constituye la principal aplicación de los MEE en el ámbito de la psicología.

Se recomienda que la fundamentación para la elección de PLS-PM frente a ACOV o GSCA, se encuentre en la coherencia entre el objetivo del estudio y el objetivo de la técnica de MEE. Se sugiere evitar otro tipo de fundamentos para dicha elección, tales como: (1). El incumplimiento de la *distribución normal de los datos*. Dado que ACOV cuenta con la posibilidad de utilizar métodos de estimación de parámetros menos restrictivos que ML, tales como ULS, ADF WLS y DWLS. Y GSCA no presenta requisitos de distribución de los datos (2). La presencia de *variables manifiestas formativas*. Ya que los resultados evidencian que ACOV además de poder realizar estimaciones de modelos que incluyen variables manifiestas formativas, realiza estimaciones con una precisión similar a PLS-PM (3). La utilización de tamaños muestrales pequeños ( $n=50$ ). Dado que las tres técnicas presentan una precisión similar en dichos tamaños muestrales, excepto cuando se utiliza un nivel de asimetría alta, instancia en que PLS-PM y GSCA son más precisas que ACOV a la hora de estimar parámetros del Modelo Estructural.

Así la utilización de un tamaño muestral pequeño y asimetría alta puede fundamentar la elección de GSCA frente a ACOV, dada la inclusión de índices de globales de bondad de ajuste, y además la obtención de estimaciones más precisas que ACOV de los parámetros del Modelo Estructural. Cabe recordar que con tamaños muestrales muy grandes ( $n=500$ ) ACOV obtiene estimaciones más precisas que PLS-PM y GSCA.

Al realizar un análisis de mediación entre variables estructurales, se sugiere utilizar ACOV, dada la tendencia de PLS-PM y GSCA a identificar efectos directos

donde no los hay y a indicar que la relación existente entre variables latentes es menos intensa de lo que realmente es.

Considerando las limitaciones de GSCA expresadas en las potentes y fundamentadas críticas realizadas por Henseler (2012), a esta técnica relativamente nueva de MEE. Se valora que el estudio de las propiedades estadísticas de las estimaciones y las características de los indicadores de bondad de ajuste de GSCA es un área necesaria de profundizar en estudio posteriores.

Así, el reconocimiento de distintas técnicas de MEE, abre el campo para el estudio del comportamiento de las mismas en distintos escenarios, debiendo tener claridad al respecto de las condiciones que fundamentan la elección de cada una de ellas al utilizarlas en ámbitos aplicados. Este trabajo de tesis doctoral pretende constituir un aporte y un paso más en dicha línea.



---

## REFERENCIAS

- Abad, F. J., Olea, J., Ponsoda, V., y García, C. (2011). *Medición en ciencias sociales y de la salud*. Madrid: Síntesis, S. A.
- Addinsoft (2009). *XLSTAT 2009*. Francia: Addinsoft. [Disponible en: <http://www.xlstat.com/en/products/xlstat-plspm/>].
- Anderson, J. C., y Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin*, 103(3), 411-423.
- Arbuckle, J. (1994). AMOS: Analysis of moment structures. *Psychometrika*, 59(1), 135-37.
- Areskoug, B. (1982). The first canonical correlation. Theoretical PLS analysis and simulation experiments. En K. G. Jöreskog y H. Wold (Eds.), *Systems under Indirect Observation*, Part II (pp. 95-117). Amsterdam: North-Holland.
- Asparouhou, T., y Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, 16(3), 397-438.
- Barroso, C., Cepeda, G., y Roldán, J. L. (2010). Applying maximum likelihood and PLS on different sample sizes: Studies on SERVQUAL model and employee behaviour model. En V. Esposito Vinzi, W. W. Chin, J. Henseler, y H. Wang (Eds.), *Handbook of Partial Least Squares. Concepts, methods and applications* (pp. 427-447). Berlin: Springer-Verlag.
- Batista-Foguet, J., y Coenders, G. (2000). *Modelos de Ecuaciones Estructurales*. Madrid: La Muralla, S. A.
- Bentler, P. M. (1995). *EQS Structural Equations Program Manual*. Encino, CA: Multivariate Software.
- Blalock, H. M. (1961). Correlation and causality: The multivariate case. *Social Forces*, 39, 246-251.
- Blalock, H. M. (1962). Four variable causal models and partial correlations. *American Journal of Sociology*, 68, 182-194.
- Blalock, H. M. (1964). *Causal inferences in nonexperimental research*. Chapel Hill: University of North Carolina Press.
- Boker, S., Neale, M., Maes, H., Wilde, M., Spiegel, M., Brick, T. R., Spies, J., Estabrook, R., Kenny, S., Bates, T. C., Mehta, P., & Fox, J. (2011). *OpenMx*:

- Multipurpose Software for Statistical Modeling*. (Version R package version 1.0.4). Virginia. [Disponible en: <http://openmx.psyc.virginia.edu>].
- Bollen, K. A. (1984). Multiple indicators: Internal consistency or no necessary relationship?. *Quality and Quantity*, 18, 377-385.
- Bollen, K. A. (1989). *Structural Equations with Latent Variables*. New York: Wiley.
- Bollen, K. A. (2002). Latent variables in psychology and the social sciences. *Annual Reviews of Psychology*, 53, 605-634.
- Bollen, K. A., Kirby, J. B., Curran, P. J., Paxton, P. M., y Chen, F. (2007). Latent variable models under misspecification. Two-stages least squares (2SLS) and Maximum likelihood (ML) estimators. *Sociological Methods and Research*, 26(1), 48-86.
- Bollen, K. A., y Long, J. S. (1993). *Testing Structural Equation Models*. Newbury Park, CA: Sage Publications.
- Boomsma, A., y Hoogland, J. J. (2001). The robustness of LISREL modeling revisited. En R. Cudeck, S. H. C. Du Toit, y D. Sörbom (Eds.), *Structural Equation Modeling: Present and Future a Festschrift in Honor of Karl Jöreskog* (pp. 139-168). Lincolnwood: Scientific Software International, Inc.
- Browne, M. W. (1984). Asymptotically distribution-free methods for the analysis of covariance structures. *British Journal of Mathematical and Statistical Psychology*, 37, 62-83.
- Browne, M. W., y Cudeck, R. (1993). Alternative ways of assessing model fit. En K. A. Bollen y J. S. Long (Eds.), *Testing Structural Equation Models* (pp. 136-162). Newbury Park, CA: Sage Publications.
- Burks, B. (1928). The relative influence of nature and nurture upon mental development: A comparative study of foster parent-foster child resemblance and true parent-true child resemblance. *The Twenty-Seventh Yearbook of the National Society for the Study of Education*, 27, 219-316.
- Castle, W. E. (1914). The nature of size factors as indicated by a study of correlation. *Carnegie Institute of Washington Publications*, 196, 51-55.
- Chin, W. W. (1995). Partial Least Squares is to LISREL as principal components analysis is to common factor analysis. *Technology Studies*, 2, 315-319.
- Chin, W. W. (1998a). The partial least squares approach to structural equation modeling. En G. A. Marcoulides (Ed.), *Modern Methods for Business Research* (pp. 295-336). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

- 
- Chin, W. W. (1998b). Issues and options on structural equation modeling. *MIS Quarterly*, 22(1), vii-xvi.
- Chin, W. W. (2001). *PLS-Graph User's Guide Version 3.0*. Houston: Soft Modeling Inc.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd edition). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Croon, M. (2002). Using predicted latent scores in general latent structure models. En G. A. Marcoulides, y I. Moustaki (Eds.), *Latent variable and latent structure models* (pp. 195-223). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- De Leeuw, J., Young, F. W., y Takane, Y. (1976). Additive structure in qualitative data: An alternating least squares method with optimal scaling features. *Psychometrika*, 41, 471-503.
- Duncan, O. D. (1975). *Introduction to Structural Equation Models*. New York: Academic Press.
- Duncan, O. D. (1966). Path analysis: Sociological examples. *American Journal of Sociology*, 72, 1-16.
- Duncan, O. D., y Hodge, R. W. (1963). Education and occupational mobility: A regression analysis. *The American Journal of Sociology*, 68, 629-644.
- Edwards, J. R., y Bagozzi, R. P. (2000). On the nature and direction of relationships between constructs and measures. *Psychological Bulletin*, 5(2), 115-174.
- Elosua, P., y Zumbo, B. D. (2008). Coeficientes de fiabilidad para escalas de respuesta categórica ordenada. *Psicothema*, 20(4), 896-901.
- Engle, R. F. (1984). Wald, likelihood ratio, and lagrange multiplier tests in econometrics. En Z. Griliches y M. D. Intriligator (Eds.), *Handbook of Econometrics*, Volume II (pp. 775-826). BV: Elsevier Science Publishers.
- Esposito Vinzi, V., Trinchera, L., y Amato, S. (2010). PLS path modeling: From foundations to recent developments and open issues for model assessment and improvement. En V. Esposito Vinzi, W. W. Chin, J. Henseler y H. Wang (Eds.), *Handbook of Partial Least Squares. Concepts, methods and applications* (pp. 47-82). New York: Springer-Verlag.
- Ferrando, P. J. (1996). Evaluación de la unidimensionalidad de los ítems mediante análisis factorial. *Psicothema*, 8(2), 397-410.
- Ferrando, P. J, y Lorenzo Seva, U. (1993). Algunas relaciones entre el modelo de un factor común y el modelo logístico de dos parámetros. *Psicothema*, 5(2), 403-412.

- Flora, D. B., y Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods*, 9(4), 466-491.
- Forero, C. G., Maydeu-Olivares, A., y Gallardo-Pujol, D. (2009). Factor analysis with ordinal indicators: A Monte Carlo study comparing DWLS and ULS estimation. *Structural Equation Modeling*, 16, 625-641.
- Fornell, C., y Bookstein, F. L. (1982). Two structural equation models: LISREL and PLS applied to consumer exit-voice theory. *Journal of Marketing Research*, 19, 440-452.
- Fox, J., Nie, Z., y Byrne, J. (2012). *sem: Structural Equation Models*. [Disponible en: <http://cran.r-project.org/web/packages/sem/index.html>].
- Fu, J. R. (2006). *VisualPLS 1.04*. [Disponible en: <http://fs.mis.kuas.edu.tw/~fred/vpls/>].
- García Ferrer, A. (1984). Procedimiento de verificación y aplicaciones para la especificación de modelos econométricos. *Estadística Española*, 102, 13-34.
- Garrido, L., Abad, F. J., y Ponsoda, V. (2011). Performance of Velicer's minimum average partial factor retention method with categorical variables. *Educational and Psychological Measurement*, 71(3), 551-570.
- Gefen, D., Rigdon, E. E., y Straub, D. (2011). An update and extension to SEM guidelines for administrative and social science research. *MIS Quarterly*, 35(2), iii-xiv.
- Gefen, D., Straub, D., y Boudreau, M. (2000). Structural equation modeling and regression: Guidelines for research practice. *Communications of the Association for Information Systems*, 4(7), 1-78.
- Gerbing, D. W., y Anderson, J. C. (1985). The effects of sampling error and model characteristics on parameter estimation for maximum likelihood confirmatory factor analysis. *Multivariate Behavioral Research*, 20, 255-271.
- Goldberger, A. S. (1971). Econometrics and psychometrics: A survey of communalities. *Psychometrika*, 36(2), 83-107.
- Golob, T. F. (2003). Structural equation modeling for travel behavior research. *Transportation Research*, 37, 1-25.
- Guo, B., Perron, B., y Gillespie, D. (2009). A systematic review of structural equation modelling in social work research. *British Journal of Social Work*, 39(8), 1556-1574.
- Haenlein, M., y Kaplan, A. (2004). A beginner's guide to partial least squares analysis. *Understanding statistics*, 3(4), 283-297.

- 
- Hair, J. F., Sarstedt, M., Ringle, C. M., y Mena, J. A. (2012). An assessment of the use of partial least squares structural equation modeling in marketing research. *Journal of the Academy of Marketing Science*, 40(3), 414-433.
- Hancock, G. R., y Mueller, R. O. (Eds.), (2006). *Structural Equation Modeling. A second course*. Connecticut: Information Age Publishing.
- Hattie, J. (1984). An empirical study of various indices for determining unidimensionality. *Multivariate Behavioral Research*, 19(1), 49-78.
- Henseler, J. (2012). Why generalized structured component analysis is not universally preferable to structural equation modeling. *Journal of Academy of Marketing Science*, 40(3), 402-413.
- Henseler, J., y Chin, W. W. (2010). A comparison of approaches for the analysis of interaction effects between latent variables using partial least squares path modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, 17(1), 82-109.
- Henseler, J., Ringle, C. M., y Sinkovics, R. R. (2009). The use of partial least squares path modeling in international marketing. *Advances in International Marketing*, 19, 1-43.
- Hooper, D., Coughlan, J., y Mullen, M. R. (2008). Structural equation modelling: Guidelines for determining model fit. *The Electronic Journal of Business Research Methods*, 6(1), 53-60. (Accedido en: Mayo, 2011), [disponible en: <http://www.ejbrm.com/volume6/issue1/p53>].
- Hoyle, R. H., y Panter, A. T. (1995). Writing about structural equation models. En R. H. Hoyle (Ed.), *Structural Equation Modeling. Concepts, Issues, and Applications* (pp. 158-176). California: Sage Publications, Inc.
- Hu, L., y Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1-55.
- Hui, B. S., y Wold, H. (1982). Consistency and consistency at large in partial least squares estimates. En K. G. Jöreskog y H. Wold (Eds.), *Systems under indirect observation: Causality, structure, prediction, Part II* (pp. 119-130). Amsterdam: North Holland.
- Hulland, J., Ryan, M. J., y Rayner, R. K. (2010). Modeling customer satisfaction: A comparative performance evaluation of covariance structure analysis versus partial least squares. En V. Esposito Vinzi, W. W. Chin, J. Henseler y H. Wang (Eds.), *Handbook of Partial Least Squares. Concepts, methods and applications* (pp. 307-325). Berlin: Springer-Verlag.
- Hwang, H. (2009). GeSCA. (Accedido en: Mayo, 2012), [disponible en <http://www.sem-gesca.org/>].

- Hwang, H. (Comunicación personal, 29 de Septiembre, 2011).
- Hwang, H., DeSarbo, W. S., y Takane, Y. (2007). Fuzzy Clusterwise Generalized Structured Component Analysis. *Psychometrika*, 72(2), 181–98.
- Hwang, H., Ho, M. H. R., y Lee, J. (2010). Generalized structured component analysis with latent interactions. *Psychometrika*, 75(2), 228-242.
- Hwang, H., Malhotra, N., Kim, Y., Tomiuk, M., y Hong, S. (2010a). A comparative study on parameter recovery of three approaches to structural equation modeling. *Journal of Marketing Research*, 47, 699-712.
- Hwang, H., Malhotra, N., Kim, Y., Tomiuk, M., y Hong, S. (2010b). Web errata for “A comparative study on parameter recovery of three approaches to structural equation modelling (Hwang, Naresh, Kim, Tomiuk, y Hong, 2010, Vol. XLVII, August, pp. 699-712)”. (Accedido en: Agosto, 2012) [disponible en: [http://www.marketingpower.com/AboutAMA/Documents/JMR\\_Web\\_Appendix/2010.4/erratum\\_comparative\\_study\\_on\\_parameter\\_recovery.pdf](http://www.marketingpower.com/AboutAMA/Documents/JMR_Web_Appendix/2010.4/erratum_comparative_study_on_parameter_recovery.pdf)].
- Hwang, H., y Takane, Y. (2004). Generalized structured component analysis. *Psychometrika*, 69(1), 81-99.
- Iacobucci, D. (2010). Structural equations modeling: Fit Indices, sample size, and advanced topics. *Journal of Consumer Psychology*, 20, 90–98.
- Jamshidian, M., y Mata, M. (2007). Advances in analysis of mean and covariance structures when data are incomplete. En S. Y. Lee (Ed.), *Handbook of Latent Variable and Related Models*, (pp. 21-44). Amsterdam: Elsevier.
- Jöreskog, K. G. (1970a). A general method for estimating a linear structural equation system. *Educational Testing Service, Research Bulletin*, 70-54.
- Jöreskog, K. G. (1970b). A general method for analysis of covariance structures. *Biometrika*, 52, 239-251.
- Jöreskog, K. G. (1971). Statistical analysis of sets of congeneric tests. *Psychometrika*, 36(2), 109-133.
- Jöreskog, K. G. (1978). Structural analysis of covariance and correlation matrices, *Psychometrika*, 43(4), 443-477.
- Jöreskog, K. G. (1979a). Analyzing psychological data by structural analysis of covariance matrix. En J. Magidson (Ed.), *Advances in Factor Analysis and Structural Equation Models* (pp. 45-100). New York: University Press.
- Jöreskog, K. G. (1979b). Structural equation models in the social sciences: Specification, Estimation and testing. En J. Magidson (Ed.), *Advances in Factor*

- 
- Analysis and Structural Equation Models* (pp. 105-127). New York: University Press.
- Jöreskog, K. G., y Goldberger, A. S. (1975). Estimation of a model with multiple indicators and multiple causes of a single latent variable. *Journal of the Statistical American Association*, 70(351), 631-639.
- Jöreskog, K. G., Gruvaeus, G. T., y van Thillo, M. (1970). ACOVS: A general computer program for the analysis of covariance structures, *Educational Testing Service, Research Bulletin*, 70-15.
- Jöreskog, K. G., y Sörbom, D. (1996-2001). *LISREL 8: User's Reference Guide*. Lincolnwood: Scientific Software International, Inc.
- Jöreskog, K. G., y Sörbom, D. (2006). *LISREL 8.8 for Windows* [Computer software]. Lincolnwood: Scientific Software International, Inc.
- Jöreskog, K. G., y van Thillo, M. (1972). LISREL: A general computer program for estimating a linear structural equation of unmeasured variables. *Educational Testing Service, Research Bulletin*, 72-56.
- Jöreskog, K. G., y Wold, H. (1982). The ML and PLS techniques for modeling with latent variables: Historical and comparative aspects. En K. G. Jöreskog y H. Wold (Eds.), *Systems under indirect observation: Causality, structure, prediction*, part I (pp.263-270). Amsterdam: North-Holland.
- Kaplan, D. (2009). *Structural Equation Modeling. Foundations and Extensions* (2nd edition). California: Sage Publications, Inc.
- Keesling, J. (1972). *Maximum likelihood approaches to causal analysis*. Ph.D. dissertation. Department of Education. New York: Wiley.
- Kim, K. H. (2005). The relation among fit indexes, power, and sample size in structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, 12(3), 368-390
- Kline, R. B. (2011). *Principles and Practice of Structural Equation Modeling* (3th edition). New York: The Guilford Press.
- Lévy, J. P., González, N., y Muñoz, M. F. (2006). Modelos de ecuaciones estructurales según el método de optimización de mínimos cuadrados parciales (PLS). En J. P. Lévy y J. Varela (Coords.), *Modelización con Estructuras de Covarianzas en Ciencias Sociales: Temas esenciales, avanzados y aportaciones especiales* (pp. 321-355). A Coruña: Netbiblo.
- Lévy, J. P., Martín, M. T., y Román, M. V. (2006). Optimización según estructuras de covarianzas. En J. P. Lévy y J. Varela (Coords.), *Modelización con Estructuras de*

- Covarianzas en Ciencias Sociales: Temas esenciales, avanzados y aportaciones especiales* (pp. 11-30). A Coruña: Netbiblo.
- Lévy, J. P., y Oubiña, J. (2006). Análisis path y modelización con variables observables. En J. P. Lévy y J. Varela (Coords.), *Modelización con Estructuras de Covarianzas en Ciencias Sociales: Temas esenciales, avanzados y aportaciones especiales* (pp. 175-191). A Coruña: Netbiblo.
- Lévy, J. P., y Varela, J. (Coords.), (2006). *Modelización con Estructuras de Covarianzas en Ciencias Sociales: Temas esenciales, avanzados y aportaciones especiales*. A Coruña: Netbiblo.
- Li, Y. (2003). *PLS-GUI User Manual – A Graphic User Interface for LVPLS (PLS-PC 1.8) – Version 1.0*. University of South Carolina, Columbia, SC.
- Loehlin, J. C. (1987). *Latent variables models. An introduction to Factor, Path and Structural Analysis*. New York: Lawrence Erlbaum Associates.
- Lohmöller, J. B. (1984). *LVPLS Program Manual*. Germany: Zentralarchiv für Empirische Sozialforschung, Universität Zu Köln.
- Lohmöller, J. B. (1987). *LVPLS Program Manual, Version 1.8*. Germany: Zentralarchiv für Empirische Sozialforschung, Universität Zu Köln.
- Lohmöller, J. B. (1989). *Latent Variable Path Modeling with Partial Least Squares*. Heidelberg: Physica-Verlag.
- Lyttkens, E. (1968). On the fixed-point property of Wold's iterative estimation method for principal components. En P. R. Krishnaiah (Ed.), *Multivariate Analysis* (pp. 335–350). New York: Academic Press.
- Lyttkens, E. (1973). The fixed-point method for estimating interdependent systems with the underlying model specification. *Journal of the Royal Statistical Society, Series A (General)*, 136, 353–394.
- MacCallum, R. C., y Browne, M. W. (1993). The use of causal indicators in covariance structure models: Some practical issues. *Psychological Bulletin*, 114(3), 533-541.
- Marcoulides, G. A. (2006). PLS: A silver bullet? *MISS Quarterly*, 30(2), iii-ix.
- Marsh, H. W., Hau, K. T., Balla, J. R., y Grayson, D. (1998). Is more ever too much? The number of indicators per factor in confirmatory factor analysis. *Multivariate Behavioral Research*, 33, 181–220.
- Marsh, H. W., Hau, K., y Grayson, D. (2005). Goodness of fit in structural equation models. En A. Maydeu-Olivares y J. J. McArdle (Eds.), *Contemporary Psychometrics* (pp. 275-340). New Jersey: Lawrence Erlbaum Associates, Inc.



- 
- Martínez Arias, R. (1995). *Psicometría: Teoría de los Tests Psicológicos y Educativos*. Madrid: Síntesis, S. A.
- Martínez-Ruiz, A., y Aluja Banet, T. (2010). PLS path modeling with mode C computational experiments. En *World Congress on Engineering (WCE) 2010*, International Association of Engineers, London, United Kingdom, 30 de Junio - 02 de Julio de 2010. (Accedido en Noviembre, 2011), [disponible en: [http://www.iaeng.org/publication/WCE2010/WCE2010\\_pp1987-1992.pdf](http://www.iaeng.org/publication/WCE2010/WCE2010_pp1987-1992.pdf)].
- McDonald, R. P. (1981). The dimensionality of tests and items. *British Journal of Mathematical and Statistical Psychology*, 34, 100-117.
- McDonald, R. P. (1996). Path analysis with composite variables. *Multivariate Behavioral Research*, 31(2), 239-270.
- Monecke, A., & Leisch, F. (2012). semPLS: Structural Equation Modeling Using Partial Least Squares. *Journal of Statistical Software*, 48(3), 1-32. [Disponible en: <http://www.jstatsoft.org/v48/i03/paper>].
- Mulaik, S. A. (2009). *Linear Causal Modeling with Structural Equations*. Boca Raton, FL: Chapman y Hall/CRC.
- Muñiz, J., y Cuesta, M. (1993). El problema de la unidimensionalidad en la medición psicológica. En M. Fornis y M. T. Anguera (Comps.), *Aportaciones recientes a la evaluación psicológica* (pp. 51-70). Barcelona: PPU, S. A.
- Muthén, B. O., du Toit, S. H. C., y Spisic, D. (1997). Robust inference using weighted least squares and quadratic estimating equations in latent variable modeling with categorical and continuous outcomes. Reporte Técnico no publicado. (Accedido en Agosto, 2011), [disponible en: <http://www.statmodel.com/wlscv.shtml>].
- Muthén, B. O., Kaplan, D., y Hollis, M. (1987). On structural equation modeling with data that are not missing completely at random. *Psychometrika*, 52(3), 431-462.
- Muthén, L., y Muthén, B. O. (1998-2010). *Mplus user's guide* (6<sup>th</sup> ed.). Los Angeles, CA: Muthén y Muthén. [Disponible en: <http://www.statmodel.com/download/usersguide/Mplus%20Users%20Guide%20v6.pdf>].
- Newsom (2005). *Some Clarifications and Recommendations on Fit Indices*. Newsom USP 655 SEM. (Accedido en: Octubre, 2011) [disponible en: [http://www.upa.pdx.edu/IOA/newsom/semclass/ho\\_fit.doc](http://www.upa.pdx.edu/IOA/newsom/semclass/ho_fit.doc)].
- Olsson, U. (1979). Maximum likelihood estimation of the polychoric correlation coefficient. *Psychometrika*, 44(4), 443-460.

- Paxton, P., Curran, P. J., Bollen, K. A., Kirby, J. B., y Chen, F. (2001). Monte Carlo experiments: Design and implementation. *Structural Equation Modeling*, 8(2), 287–312.
- R Development Core Team (2011). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>.
- Raykov, T., Marcoulides, G. A., y Boyd, J. (2003). Using SEM programs to perform matrix manipulations and data simulation. *Structural Equation Modeling*, 10(2), 312-322.
- Reinartz, W., Haenlein, M., y Henseler, J. (2009). An empirical comparison of the efficacy of covariance-based and variance-based SEM. *International Journal of Research in Marketing*, 26, 332-344.
- Rigdon, E. E. (1998). Structural Equation Modeling. En G. A. Marcoulides (Ed.), *Modern Methods for Business Research* (pp. 251-294). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Rigdon, E. E., Ringle, C. M., y Sarstedt, M. (2010). Structural modeling of heterogeneous data with partial least squares. *Review of Marketing Research*, 7, 255-296.
- Ringle, C. M., Götz, O., Wetzels, M., y Wilson, B. (2009). On the Use of Formative Measurement Specifications in Structural Equation Modeling: A Monte Carlo Simulation Study to Compare Covariance-Based and Partial Least Squares Model Estimation Methodologies. *MPRA*, paper No 15390. [Disponible en: <http://mpa.ub.uni-muenchen.de/15390/>].
- Ringle, C. M., Sarstedt, M., y Straub, D. W. (2012). A Critical Look at the Use of PLS-SEM in *MIS Quarterly*. *MIS Quarterly*, 36(1), iii-xiv.
- Ringle, C. M, Wende, S., y Will, A. (2005a). *SmartPLS – Version 2.0*. Hamburg: University of Hamburg.
- Ringle, C. M, Wende, S., y Will, A. (2005b). Customer segmentation with FIMIX-PLS. En T. Aluja, J. Casanovas, V. Esposito Vinzi, A. Morineau, y M. Tenenhaus (Eds.), *Proceedings of PLS-05 International Symposium* (pp. 507-514). Paris: SPAD Test and Go.
- Ringle, C. M, Wende, S., y Will, A. (2010). Finite mixture partial least squares analysis: Methodology and numerical examples. En V. Esposito Vinzi, W. W. Chin, J. Henseler, y H. Wang (Eds.), *Handbook of Partial Least Squares. Concepts, methods and applications* (pp. 195-218). Berlin: Springer-Verlag.

- 
- Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2), 1-36. [Disponible en <http://www.jstatsoft.org/v48/i02/paper/>].
- Ruiz, M. A. (2000). *Introducción a los Modelos de Ecuaciones Estructurales*. Madrid: UNED.
- Ruiz, M. A., Pardo, A., y San Martín, R. (2010). Modelos de Ecuaciones Estructurales. *Papeles del Psicólogo*, 31(1), 34-45. (Accedido en: Enero, 2011), [disponible en: <http://redalyc.uaemex.mx/pdf/778/77812441004.pdf>].
- Sánchez, G., y Trinchera, L. (2012). *plspm: Partial Least Squares Data Analysis Method*. (Accedido en: Abril, 2012), [disponible en: <http://cran.r-project.org/web/packages/plspm/index.html>].
- SAS Institute Inc. (2008). *SAS/STAT® 9.2 User's Guide*. Cary, NC: SAS Institute Inc.
- Schermelleh-Engel, K., Moosbrugger, H., y Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of psychological research online*, 8, 23-74.
- Schreiber, J. B., Stage, F. K., King, J., Nora, A., y Barlow, E. A. (2006). Reporting structural equation modeling and confirmatory factor analysis results. A review. *The Journal of Educational Research*, 99(6), 323-337.
- Sellin, N. (1989). *PLSPATH - Version 3.01. Application Manual*. Universität Hamburg, Hamburg.
- Skrondal, A., y Laake, P. (2001). Regression among factor scores. *Psychometrika*, 66(4), 563-576.
- Steiger, J. H. (1988). Aspects of person-machine communication in structural modeling of correlations and covariances. *Multivariate Behavioral Research*, 23, 281-290.
- Steiger, J. H. (2007). Understanding the limitations of global fit assessment in structural equation modeling. *Personality and Individual Differences*, 42(5), 893-898.
- Tenenhaus, M. (2008). Component-based structural equation modelling. *Total Quality Management and Business Excellence*, 19(7), 871-886.
- Tenenhaus, M., Amato, S., y Esposito Vinzi, V. (2004). A global goodness-of-fit index for PLS structural equation modelling. *Proceedings of the XLII SIS Scientific Meeting*, Vol I. Contributed Papers CLEUP, Padova, pp. 739-742.
- Tenenhaus, M., Esposito Vinzi, V., Chatelin, Y., y Lauro, C. (2005). PLS path modeling. *Computational Statistics and Data Analysis*, 48, 159-205.

- Tenenhaus, M., Mauger, E., y Guinot, C. (2010). Use of ULS-SEM and PLS-SEM to measure a group effect in a regression model relating two blocks of binary variables. En V. Esposito Vinzi, W. W. Chin, J. Henseler y H. Wang (Eds.), *Handbook of Partial Least Squares. Concepts, methods and applications* (pp. 125-140). Berlin: Springer-Verlag.
- Test and Go (2006). *Spad Version 6.0.0*. Paris, France.
- Urbach, N. y Ahlemann, F. (2010). Structural equation modeling in information systems research using partial least squares. *Journal of Information Technology Theory and Application (JITTA)*, 11(2), 5-40.
- Vilares, M. J., Almeida, M. H., y Coehlo, P. S. (2010). Comparison of likelihood and PLS estimators for structural equation modeling: A simulation with consumer satisfaction data. En V. Esposito Vinzi, W. W. Chin, J. Henseler y H. Wang (Eds.), *Handbook of Partial Least Squares. Concepts, methods and applications* (pp. 289-305). Berlin: Springer-Verlag.
- Werts, C., y Linn, R. (1970). Path analysis: Psychological examples. *Psychological Bulletin*, 74, 193-212. (Accedido en: Enero, 2011), [disponible en: <http://psycnet.apa.org/journals/bul/74/3/193/>].
- Wiley, D. (1973). The identification problem for structural equation models with unmeasured variables. En A. S. Goldberger y O. D. Duncan (Eds.), *Structural Equation Models in the Social Sciences* (pp. 69-83). New York: Seminar Press.
- Williams, L. J., Edwards, J. R., y Vandenberg, R. J. (2003). Recent advances in causal modeling methods for organizational and management research. *Journal of Management*, 29(6), 903-936.
- Willmott, C. J., y Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, 30, 79-82.
- Wold, H. (1954). Causality and econometrics. *Econometrica*, 22(2), 162-177.
- Wold, H. (1965). A fixed-point theorem with econometric background. Part I: The theorem. *Arkiv for Matematik*, 6, 209-240.
- Wold, H. (1966). Estimation of principal components and related models by iterative least squares. En P. R. Krishnaiah (Ed.), *Multivariate Analysis*, Vol I (pp. 391-420). New York: Academic Press.
- Wold, H. (1973). Non-linear iterative partial least squares (NIPALS) modeling: Some current development. En P. R. Krishnaiah (Ed.), *Multivariate Analysis*, Vol III (pp. 383-407). New York: Academic Press.
- Wold, H. (1974). Causal flows with latent variables: Partings of the ways in the light of NIPALS modelling. *European Economic Review*, 5, 67-86.

- 
- Wold, H. (1975a). Path models with latent variables: the NIPALS approach. En H. M. Blalock, A. Aganbegian, F. M. Borodkin, R. Boudon y V. Capecchi (Eds.), *Quantitative Sociology: International Perspectives on Mathematical and Statistical Modeling* (pp. 307-357). New York: Academic Press, Inc.
- Wold, H. (1975b). Soft modelling by latent variables in the non-iterative partial least squares (NIPALS) approach. En J. Gani (Ed.), *Perspectives in Probability and Statistics, Papers in Honour of M.S. Barlett* (pp. 117-142). London: Academic Press.
- Wold, H. (1977). On the transition from pattern cognition to model building. En R. Henn y O. Moeschlin (Eds.), *Mathematical Economics and Game Theory: Essays in Honor of Oskar Morgenstern* (pp. 536-549). Berlin: Springer-Verlag.
- Wold, H. (1980). Model construction and evaluation when theoretical knowledge is scarce: Theory and application of partial least squares. En J. Kmenta y J. Ramsey (Eds.), *Evaluation of Econometric Models* (pp. 47-74). New York: Academic Press.
- Wold, H. (1981). *The Fixed Point Approach to Interdependent Systems*. Amsterdam: North Holland.
- Wold, H. (1982a). Soft modeling: The basic design and some extensions. En K. G. Jöreskog y H. Wold (Eds.), *Systems under Indirect Observation, Part II* (pp. 1-54). Amsterdam: North-Holland.
- Wold, H. (1982b). Systems under indirect observation using PLS. En C. Fornell (Ed.), *A Second Generation of Multivariate Analysis, Volume 1* (pp. 325-347). New York: Praeger Publishers.
- Wold, H. (1985). Partial least squares. En S. Kotz y N. L. Johnson (Eds.), *Encyclopedia of Statistical Sciences*, Vol. 6 (pp. 581-591). New York: Wiley.
- Wold, H., y Lyttkens, E. (1969). Nonlinear iterative partial least squares (NIPALS) estimation procedures. *Bulletin of the International Statistics Institute*, 43, 29-51.
- Wright, S. (1918). On the nature of size factors. *Genetics*, 3, 367-374.
- Wright, S. (1920). The relative importance of heredity and environment in determining the piebald pattern of guinea-pigs. *Proceedings of the National Academy of Sciences*, 6, 320-332.
- Wright, S. (1921). Correlation and causation. *Journal of Agricultural Research*, 20, 557-585.
- Wright, S. (1923). The theory of path coefficients: A reply to Niles's criticism. *Genetics*, 8, 239-255.
- Wright, S. (1932). General, group and special size factors. *Genetics*, 17, 603-619.

Wright, S. (1934). The method of path coefficients. *Annals of Mathematical Statistics*, 5, 161-215.

# ANEXOS





---

**Anexo A.** Descripción de la sintaxis para la simulación de datos con escala ordinal con Matlab, a partir de la matriz de varianzas y covarianzas del modelo con variables manifiestas reflectivas, modelo más especificado.

- 1.- Se borran registros del espacio de trabajo, se define un valor de semilla para asegurar que los datos generados sean siempre los mismos.

```
clear
semilla=300000;
randn('state',semilla);
```

- 2.- Se definen valores para: el número de muestras, número de sujetos, número de factores, número de ítems por factor, y el número total de ítems.

```
nmuestras=500; nsuj=500; nfact=3; nitemsfact=3; Nitems=nfact*nitemsfact;
```

- 3.- Se definen los pesos que tendrán las variables manifiestas de cada factor.

```
cargasf1=0.7; cargasf2=0.7; cargasf3=0.7;
```

- 4.- Se construye la matriz lambda para las variables manifiestas exógenas y para las variables manifiestas endógenas por separado y una matriz Lambda con las cargas factoriales de ambos modelos de medida.

```
a1=ones(1,nitemsfact); a2=a1*cargasf1; a3=zeros(nitemsfact,1);
a4=[a2',a3,a3];
b1=a1*cargasf2; b2=horzcat(b1',a3);
c1=a1*cargasf3; c2=horzcat(a3,c1');
lambdax=a2';
lambday=vertcat(b2,c2);
e1=zeros(nitemsfact*2,1); e2=[e1,lambday];
Lambda=vertcat(a4,e2);
```

- 5.- Se define Phi, varianza de la variable latente exógena.

```
Phi=1;
```

- 6.- Se define la matriz Gamma, de coeficientes de regresión de la variable exógena sobre las variables endógenas.

```
Gamma=[0.6;0];
```

- 7.- Se define la matriz Identidad ( $I$ ) y la matriz Beta, de coeficientes de regresión de la variable latente endógena sobre la otra variable latente endógena.

```
I=eye(2);
Beta=[0 0;0.6 0];
```

- 8.- Se define la matriz Psi, calculando las varianzas de las Zetas necesarias para que las Etas tengan varianzas igual a 1.

```
P=1-(Gamma.^2);
PP=[P(1,1) P(1,1)];
Psi=diag(PP);
```

- 9.- Se definen las matrices ( $SIGMA_{xy}$ ,  $SIGMA_{yx}$ ,  $SIGMA_{xx}$  y  $SIGMA_{yy}$ ) que componen la matriz reproducida ( $SIGMA$ ) del modelo.

---

```

SIGMAxy=lambdax*Phi*Gamma'*( (I-Beta)')^-1*lambday';
SIGMAyx=lambday*( (I-Beta))^-1*Gamma*Phi*lambdax';

SIGMAx=lambdax*Phi*lambdax';
Td=eye(nitemsfact);
Te=eye(nitemsfact*2);
d=diag(SIGMAx);dd=diag(d);
Thetadelta=Td-dd;
SIGMAxx=SIGMAx+Thetadelta;

SIGMAy=lambday*( (I-Beta)^-1*(Gamma*Phi*Gamma'+Psi))*((I-Beta)')^-1*lambday';
ddd=diag(SIGMAy);dddd=diag(ddd);
Thetaepsilon=Te-dddd;
SIGMAyy=SIGMAy+Thetaepsilon;

SIGMA=horzcat(vertcat(SIGMAyy,SIGMAxy),vertcat(SIGMAyx,SIGMAxx));

```

- 10.- Se construye una matriz tridimensional (D3) para guardar los datos generados, con x=cantidad de sujetos, y=cantidad de ítems; y z=cantidad de muestras.

```
D3=zeros(nsuj,Nitems,nmuestras);
```

- 11.- Se inicia un ciclo “for” en el cual se determina que la cantidad de ciclos será igual a la cantidad de muestras definida en el paso 2.

```
for i=1:nmuestras
```

- 12.- Se realiza la descomposición de Cholesky de la matriz reproducida SIGMA, obteniendo la matriz triangular U.

```
U=chol(SIGMA);
```

- 13.- Se genera una matriz (z), de tamaño igual al número de sujetos (filas) por número de ítems (columnas), de números aleatorios de una distribución normal.

```
z=randn(nsuj,Nitems);
```

- 14.- Se multiplica la matriz (z) de números aleatorios con distribución normal por la matriz U. Obteniéndose una matriz (x) de variables continuas muestrales.

```
x=z*U;
```

- 15.- Se genera un histograma con ajuste a la curva normal de la variable 1 de la matriz (x).

```
subplot(1,5,[1 2])
histfit(x(:,1))
```

- 16.- Se inicia un ciclo “for” en el cual se categorizan los datos continuos en 5 puntos, de acuerdo a los umbrales señalados por Garrido, Abad, Ponsoda (2011).

```
for j=1:Nitems
    for h=1:nsuj
        if x(h,j)<=-1.8
            x(h,j)=1;
        elseif -1.8<x(h,j) && x(h,j)<=-0.6
            x(h,j)=2;
        elseif -0.6<x(h,j) && x(h,j)<=0.6
```

```

        x(h,j)=3;
    elseif 0.6<x(h,j) && x(h,j)<=1.8
        x(h,j)=4;
    else
        x(h,j)=5;
    end
end
end
end

```

17.- Se guardan los datos en la matriz tridimensional D3.

```
D3(:,:,i)=x;
```

18.- Se calculan la asimetría y curtosis de las variables manifiestas para cada muestra.

```
Skew(i,:)=skewness(D3(:,:,i));
Kurt(i,:)=kurtosis(D3(:,:,i));
```

19.- Se exportan los datos a una planilla Excel, con el nombre “datos\_500s\_a0\_matco” en la cual, cada hoja corresponde a una muestra de datos ordinales con la estructura determinada por la matriz SIGMA y se finaliza el ciclo “for” iniciado en el punto 11.

```
xlswrite('datos_500s_a0_matco',D3(:,:,i),i)
end
```

20.- Se genera un histograma con 5 puntos en el eje x, con ajuste a la curva normal de la variable 1, de la muestra 1, de la matriz D3.

```
subplot(1,5,[4 5])
histfit(D3(:,1,1),5)
```

Sintaxis Matlab:

```

clear
semilla=300000;
randn('state',semilla);
nmuestras=500; nsuj=500; nfact=3; nitemsfact=3; Nitems=nfact*nitemsfact;
cargasf1=0.7; cargasf2=0.7; cargasf3=0.7;
a1=ones(1,nitemsfact); a2=a1*cargasf1; a3=zeros(nitemsfact,1); a4=[a2',a3,a3];
b1=a1*cargasf2; b2=horzcat(b1',a3);
c1=a1*cargasf3; c2=horzcat(a3,c1');
lambdax=a2';
lambday=vertcat(b2,c2);
e1=zeros(nitemsfact*2,1); e2=[e1,lambday];
Lambda=vertcat(a4,e2);
Phi=1;
Gamma=[0.6;0];
I=eye(2);
Beta=[0 0;0.6 0];
P=1-(Gamma.^2);
PP=[P(1,1) P(1,1)];
Psi=diag(PP);
SIGMAxy=lambdax*Phi*Gamma'*((I-Beta)')^-1*lambday';
SIGMAyx=lambday*((I-Beta))^-1*Gamma*Phi*lambdax';
SIGMAx=lambdax*Phi*lambdax';
Td=eye(nitemsfact);
Te=eye(nitemsfact*2);
d=diag(SIGMAx); dd=diag(d);
Thetadelta=Td-dd;
SIGMAxx=SIGMAx+Thetadelta;
SIGMAy=lambday*((I-Beta)^-1*(Gamma*Phi*Gamma'+Psi))*((I-Beta)')^-1*lambday';

```

---

```
ddd=diag(SIGMAy);dddd=diag(ddd);
Thetaepsilon=Te-dddd;
SIGMAyy=SIGMAy+Thetaepsilon;
SIGMA=horzcat(vertcat(SIGMAyy,SIGMAxy),vertcat(SIGMAyx,SIGMAxx));
D3=zeros(nsuj,Nitems,nmuestras);
Skew=zeros(nmuestras,Nitems);
Kurt=zeros(nmuestras,Nitems);
for i=1:nmuestras
    U=chol(SIGMA);
    z=randn(nsuj,Nitems);
    x=z*U;
    subplot(1,5,[1 2])
    histfit(x(:,1))
    for j=1:Nitems
        for h=1:nsuj
            if x(h,j)<=-1.8
                x(h,j)=1;
            elseif -1.8<x(h,j) && x(h,j)<=-0.6
                x(h,j)=2;
            elseif -0.6<x(h,j) && x(h,j)<=0.6
                x(h,j)=3;
            elseif 0.6<x(h,j) && x(h,j)<=1.8
                x(h,j)=4;
            else
                x(h,j)=5;
            end
        end
    end
    D3(:, :, i)=x;
    Skew(i,:)=skewness(D3(:, :, i));
    Kurt(i,:)=kurtosis(D3(:, :, i));
    xlswrite('datos_500s_a0',D3(:, :, i),i)
end
subplot(1,5,[4 5])
histfit(D3(:,1,1),5)
```

---

**Anexo B.** Matriz de varianzas y covarianzas poblacional,  $\Sigma(\theta)$ , del modelo con variables manifiestas reflectivas, más especificado.

---

	Y1	Y2	Y3	Y4	Y5	Y6	X1	X2	X3
Y1	1								
Y2	0,490	1							
Y3	0,490	0,490	1						
Y4	0,294	0,294	0,294	1					
Y5	0,294	0,294	0,294	0,490	1				
Y6	0,294	0,294	0,294	0,490	0,490	1			
X1	0,294	0,294	0,294	0,176	0,176	0,176	1		
X2	0,294	0,294	0,294	0,176	0,176	0,176	0,490	1	
X3	0,294	0,294	0,294	0,176	0,176	0,176	0,490	0,490	1

---

**Anexo C.** Descripción de la sintaxis PRELIS para el cálculo de la matriz de correlaciones policóricas.

1.- Se establece el título (TI) de la sintaxis.

```
TI Corr_Polic
```

2.- Se define los datos (DA), en cuanto al número de variables manifiestas del modelo (NI).

```
DA NI=9
```

3.- Se determina el nombre de las variables manifiestas (LA), comenzando por las variables manifiestas asociadas a las variables latentes endógenas.

```
LA  
1 2 3 4 5 6 7 8 9
```

4.- Se determina la ruta para acceder la matriz de datos brutos (RA).

```
RA='C:\automatizac\ç.txt'
```

5.- Se define que la matriz utilizada es la matriz de correlaciones policóricas (MA) y se define el nombre del archivo en que se guardará la matriz de correlaciones policóricas (SM).

```
OU MA=PM SM=ç.PML
```

**Sintaxis PRELIS:**

```
TI Corr_Polic  
DA NI=9  
LA  
1 2 3 4 5 6 7 8 9  
RA='C:\automatizac\ç.txt'  
OU MA=PM SM=ç.PML
```

---

**Anexo D.** Descripción de la sintaxis LISREL para el análisis ACOV del modelo con variables manifiestas reflectivas, más especificado.

1.- Se establece el título (TI) de la sintaxis.

```
TI 1Exo_2Endo_ord_policorica_ULS
```

2.- Se define los datos (DA), en cuanto al número de variables manifiestas del modelo (NI), el tamaño de la muestra (NO) y que la matriz analizada (MA) será la de correlaciones policóricas (PM).

```
DA NI=9 NO=500 MA=PM
```

3.- Se determina la ruta (FI) para acceder la matriz de correlaciones policóricas (PM).

```
PM FI=ç.PML
```

4.- Se determina el orden en que serán ingresadas las variables al modelo (SE), comenzando por las variables manifiestas asociadas a las variables latentes endógenas.

```
SE
1 2 3 4 5 6 7 8 9 /
```

5.- Se especifica el modelo (MO), determinando la cantidad de variables manifiestas exógenas (NX), la cantidad de variables manifiestas endógenas (NY), la cantidad de variables latentes exógenas (NK) y la cantidad de variables latentes endógenas (NE). Luego se especifica que\* Beta (BE) es una matriz completa (FU), y fija (FI), que la matriz Gamma (GA) es completa (FU) y libre (FR), que la matriz Phi (PH) es asimétrica (SY) y libre (FR), matrices Psi (PS), Theta Delta (TD) y Theta Epsilon (TE) son diagonales (DI) y libres (FR) para la estimación de sus parámetros.

```
MO NX=3 NY=6 NK=1 NE=2 BE=FU,FI GA=FU,FR PH=SY,FR PS=DI,FR TD=DI,FR
TE=DI,FR
```

6.- Se definen los nombres de las variables latentes endógenas (LE) y la variable latente exógena (LK).

```
LE
ETA1 ETA2
LK
KSI
```

7.- Se fija a 1 los valores (VA) de Lambda, correspondientes a la primera variable manifiesta asociada a cada variable latente, para dar métrica a las mismas.

```
VA 1 LY(1,1) LY(4,2) LX(1,1)
```

8.- Se definen los parámetros libres (FR) de Lambda que serán estimados.

```
FR LY(2,1) LY(3,1) LY(5,2) LY(6,2) LX(2,1) LX(3,1)
```

---

\* En los análisis se especificó: BE=FU PS=SY TE=SY TD=SY. Frente a lo cual LISREL asume las especificaciones por defecto señaladas en el punto 5, las cuales junto con las especificaciones de parámetros realizadas en los puntos 7, 8, 9 y 10, constituyen la especificación del modelo, que conlleva la estimación de los parámetros requeridos por el investigador.

9.- Se define el parámetro fijo (FI) de Gamma (GA) a 0.

```
FI GA(2,1)
```

10.- Se define el parámetro libre de Beta.

```
FR BE(2,1)
```

11.- Opcionalmente, se puede definir la representación del diagrama estructural.

```
PD
```

12.- Se define que para la obtención de resultados (OU) el método de estimación de parámetros (ME) será Unweighted Least Squares (UL), se solicita la solución estandarizada (SS) y se define el número máximo de iteraciones (AD) para comprobar la admisibilidad de la solución.

```
OU ME=UL SS AD=100
```

### Sintaxis LISREL:

```
TI 1Exo_2Endo_ord_policorica_ULS
DA NI=9 NO=500 MA=PM
PM FI=ç.PML
SE
1 2 3 4 5 6 7 8 9 /
MO NX=3 NY=6 NK=1 NE=2 BE=FU,FI GA=FU,FR PH=SY,FR PS=DI,FR TD=DI,FR TE=DI,FR
LE
ETA1 ETA2
LK
KSI
VA 1 LY(1,1) LY(4,2) LX(1,1)
FR LY(2,1) LY(3,1) LY(5,2) LY(6,2) LX(2,1) LX(3,1)
FI GA(2,1)
FR BE(2,1)
OU ME=UL SS AD=100
```



---

**Anexo E.** Descripción de la sintaxis R para el análisis PLS-PM del modelo con variables manifiestas reflectivas, más especificado.

1.- Se abre el paquete “plspm”.

```
library(plspm)
```

2.- Se asignan los datos al objeto denominado “A”.

```
A <-read.table("C:\\Automatizac\\ç.txt")
```

3.- Se especifica la matriz Beta y se asigna al objeto denominado “inner”.

```
inner <- matrix(c(0,0,0,1,0,0,0,1,0),3,3,byrow=TRUE)
```

4.- Se definen la matriz Lambda y se asocia al objeto “outer”, comenzando por las variables manifiestas asociadas a la variable latente exógena.

```
outer <- list(c(7,8,9),c(1,2,3),c(4,5,6))
```

5.- Se define la modalidad refleja (“A”) para los tres bloques de variables manifiestas de cada componente y se asignan al objeto denominado “mod”.

```
mod <- c("A","A","A")
```

6.- Se realiza el análisis PLS-PM utilizando los objetos definidos en los pasos 2, 3, 4 y 5, determinando el esquema Path weighting para la estimación de parámetros; y se asigna al objeto denominado “C”.

```
C <-plspm(A,inner=inner,outer=outer,modes=mod,scheme="path")
```

7.- Se define la ruta del archivo en que serán guardados los resultados.

```
sink(file="C:\\Automatizac\\ç.txt")
```

8.- Se solicita el resumen del análisis PLS-PM del objeto “C”

```
summary(C)
```

9.- Se cierra la función que guarda los archivos especificados en el paso anterior.

```
sink()
```

**Sintaxis R:**

```
library(plspm)
A <-read.table("C:\\Automatizac\\ç.txt")
inner <- matrix(c(0,0,0,1,0,0,0,1,0),3,3,byrow=TRUE)
outer <- list(c(7,8,9),c(1,2,3),c(4,5,6))
mod <- c("A","A","A")
C <-plspm(A,inner=inner,outer=outer,modes=mod,scheme="path")
sink(file="C:\\Automatizac\\ç.txt")
summary(C)
sink()
```

**Anexo F.** Descripción de la sintaxis Visual FoxPro, para la ejecución automatizada del análisis ACOV con LISREL, del modelo con variables manifiestas reflectivas, más especificado.

- 1.- Se borra la ventana de salida y se cierran archivos abiertos.

```
CLEAR  
CLOSE ALL
```

- 2.- Se desactiva la solicitud de confirmación para modificar archivos.

```
SET SAFETY OFF
```

- 3.- Se establece la ruta del directorio de trabajo.

```
SET PATH TO c:\automatizac
```

- 4.- Se inicia un ciclo “for”, determinándose la cantidad de ciclos a realizar.

```
FOR i = 1 TO 500
```

- 5.- Se crea la variable “h1”, la cual tomará los valores de “i” definidos en el punto 4.

```
h1 = LTRIM(STR(i))
```

- 6.- Se agrega la palabra “Hoja” a los valores que tomará la variable “h1”.

```
h1= 'Hoja'+h1
```

- 7.- Se crea una tabla denominada “lisrel” con 80 carctéres.

```
CREATE TABLE lisrel ( linea c(80))
```

- 8.- Se especifica el nombre del archivo del cual se extraen la sintaxis LISREL, la cual es añadida a la tabla “lisrel” creada en el paso anterior.

```
APPEND FROM 1Exo_2Endo_ord_policorica_ULS.LS8 TYPE sdf
```

- 9.- Se posiciona el control de ejecución en la línea nº 3 de la Tabla con la sintaxis de LISREL (en la que se especifica la ruta para acceder a la matriz de correlaciones policóricas, con el nombre de archivo definido por “ç”).

```
GOTO 3
```

- 10.- Se crea la variable “h3”, que contiene la ruta para acceder al archivo que contiene la matriz de correlaciones policóricas que será analizada.

```
h3= 'resprelis\' +h1
```

- 11.- Se crea la variable “instrucc2”, que constituye una línea de la tabla “lisrel”.

```
instrucc2 = lisrel.linea
```

- 12.- Se cambia el nombre del archivo que contiene los datos (“ç”) por los valores de la variable “h3”.

```
instrucc2=STRTRAN(instrucc,'ç',h3)
```

- 13.- Se reemplaza el contenido de la línea nº 3 del archivo de sintaxis LISREL por el contenido de "instrucc" (con el nombre de los archivos de datos).

```
replace lisrel.linea WITH instrucc
```

- 14.- Se crea la variable "l1", que recibe la ruta con el nombre del archivo LISREL que tiene la línea nº 3 reemplazada.

```
l1 ='c:\automatizac\ctrlisrel\'+h1+'.lpj'
```

- 15.- Se graba el archivo LISREL con el nombre archivo definido en el paso 14.

```
COPY TO &l1 TYPE sdf
```

- 16.- Se crea la variable "a", que recibe la ruta con el nombre del programa batch de LISREL.

```
a='C:\Archivos de programa\LISREL88\lisrel88.EXE'
```

- 17.- Se crea la variable "r1", que recibe la ruta con el nombre del archivo para guardar resultados.

```
r1 ='c:\automatizac\reslisrel\'+h1+'.out'
```

- 18.- Se ejecuta LISREL con la sintaxis definida por "l1" y guardando los resultados en "r1".

```
RUN/N&a &l1 &r1
```

- 19.- Se define el tiempo de pausa para continuar con el siguiente ciclo.

```
WAIT TIMEOUT 0.001
```

- 20.- Se muestra en pantalla el valor de "i", correspondiente al ciclo "for" que se está ejecutando.

```
?i
```

- 21.- Se finaliza el ciclo "for".

```
ENDFOR
```

### Sintaxis Visual FoxPro:

```
CLEAR
CLOSE ALL
SET SAFETY OFF
SET PATH TO c:\automatizac

FOR i = 1 TO 500

h1 = LTRIM(STR(i))
h1= 'Hoja'+h1

CREATE TABLE lisrel ( linea c(80))
APPEND FROM 1Exo_2Endo_ord_policorica_ULS.LS8 TYPE sdf

GOTO 3
```

## Anexos

---

```
h3= 'resprelis\' +h1
instrucc2 = lisrel.linea
instrucc2=STRTRAN(instrucc2,'ç',h3)
replace lisrel.linea WITH instrucc2

l1 ='c:\automatizac\ctrlisrel\' +h1+'.lpj'

COPY TO &l1 TYPE sdf

a="'C:\Archivos de programa\LISREL88\lisrel88.EXE"'

r1 ='c:\automatizac\reslisrel\' +h1+'.out'

RUN/N&a &l1 &r1

WAIT TIMEOUT 0.001
?i
ENDFOR
```

---

**Anexo G.** Descripción de la sintaxis Visual FoxPro, para la ejecución automatizada del análisis PLS-PM con R, del modelo con variables manifiestas reflectivas, más especificado.

1.- Se borra la ventana de salida y se cierran archivos abiertos.

```
CLEAR
CLOSE ALL
```

2.- Se desactiva la solicitud de confirmación para modificar archivos.

```
SET SAFETY OFF
```

3.- Se establece la ruta del directorio de trabajo.

```
SET PATH TO c:\automatizac
```

4.- Se inicia un ciclo “for”, determinándose la cantidad de ciclos a realizar.

```
FOR i = 1 TO 500
```

5.- Se crea la variable “h1”, la cual tomará los valores de “i” definidos en el punto 4.

```
h1 = LTRIM(STR(i))
```

6.- Se agrega la palabra “Hoja” a los valores que tomará la variable “h1”.

```
h1= 'Hoja'+h1
```

7.- Se crea una tabla Excel con 9 columnas de 5 caracteres y tantas filas como ciclos se hayan definido.

```
CREATE TABLE excel (a i(5), b i(5), c i(5), d i(5), e i(5), f i(5), g i(5),
h i(5), i i(5) )
```

8.- Se especifica el nombre del archivo del cual se extraen los datos para ser añadidos a la tabla Excel creada en el paso anterior. Sustituyendo el nombre de cada hoja por el valor de “h1”.

```
APPEND FROM datos_500s_a0_matco.xls TYPE xl8 SHEET &h1
```

9.- Se crea la variable “d1”, que recibe el nombre de los archivos de datos con formato de texto.

```
d1 ='c:\automatizac\datos\' +h1+'.txt'
```

10.- Se copian los datos en archivos de texto creados en paso 9.

```
COPY TO &d1 TYPE sdf
```

11.- Se crea una tabla denominada “R” con 85 caracteres, que recibirá la sintaxis de R.

```
CREATE TABLE R ( linea c(85))
```

12.- Se añade la sintaxis R, contenida en el archive especificado, a la tabla Excel creada en el paso 11.

---

```
APPEND FROM plspm_ref_w TYPE sdf
```

- 13.- Se posiciona el control de ejecución en la línea nº 2 de la Tabla con la sintaxis de R (en la que se especifica la ruta para acceder a los datos brutos, el nombre de archivo es definido por “ç”).

```
GOTO 2
```

- 14.- Se crea la variable “h2”, con el contenido especificado (parte de la ruta para acceder a los datos).

```
h2= 'datos\\'+h1
```

- 15.- Se crea la variable “instrucc”, que contendrá una línea de comandos de R.

```
instrucc = R.linea
```

- 16.- Se cambia el nombre del archivo que contiene los datos (“ç”) por el contenido de la variable “h2”.

```
instrucc=STRTRAN(instrucc,'ç',h2)
```

- 17.- Se reemplaza el contenido de la línea nº 2 del archivo de sintaxis R por el contenido de “instrucc” (con el nombre del archivo de datos).

```
replace R.linea WITH instrucc
```

- 18.- Se posiciona el control de ejecución en la línea nº 7 de la Tabla con la sintaxis de R (en la que se especifica la ruta para guardar los resultados, el nombre de archivo es definido por “ç”).

```
GOTO 7
```

- 19.- Se crea la variable “h3”, con el contenido especificado.

```
h3= 'resR\\'+h1
```

- 20.- Se crea la variable “instrucc2”, que recibe el contenido de la línea nº 7.

```
instrucc2 = R.linea
```

- 21.- Se cambia el nombre del archivo que en el que se guardarán los resultados (“ç”) por el contenido de la variable “h3”.

```
instrucc2=STRTRAN(instrucc2,'ç',h3)
```

- 22.- Se reemplaza el contenido de la línea nº 7 del archivo de sintaxis R por el contenido de “instrucc2” (con el nombre y ruta del archivo de resultados).

```
replace R.linea WITH instrucc2
```

- 23.- Se crea la variable “l1”, que recibe la ruta con el nombre del archivo R que tiene las líneas 2 y 7 reemplazadas.

```
l1 ='c:\automatizac\ctrR\' +h1+'.r'
```

24.- Se graba el archivo R con el nombre archivo definido en el paso 23 en la variable "i1".

```
COPY TO &i1 TYPE sdf
```

25.- Se crea la variable "a", que recibe la ruta con el nombre del programa batch de R.

```
a="C:\Archivos de programa\R\R-2.14.0\bin\i386\Rscript.exe"
```

26.- Se crea la variable "r1", que recibe la ruta con el nombre del archivo para guardar resultados.

```
r1 = 'c:\automatizac\resR\' + h1 + '.txt'
```

27.- Se ejecuta R con la sintaxis definida por "i1" y guardando los resultados en "r1".

```
RUN/N&a &i1 &r1
```

28.- Se define el tiempo de pausa para continuar con el siguiente ciclo.

```
WAIT TIMEOUT 0.80
```

29.- Se muestra en pantalla el valor de "i", correspondiente al ciclo "for" que se está ejecutando.

```
?i
```

30.- Se finaliza el ciclo "for".

```
ENDFOR
```

Sintaxis Visual FoxPro:

```
CLEAR
CLOSE ALL
SET SAFETY OFF
SET PATH TO c:\automatizac

FOR i = 1 TO 500

h1 = LTRIM(STR(i))
h1= 'Hoja'+h1

CREATE TABLE excel (a i(5), b i(5), c i(5), d i(5), e i(5), f i(5), g i(5), h
i(5), i i(5) )

APPEND FROM datos_500s_a0_matco.xls TYPE XL8 SHEET &h1

d1 ='c:\automatizac\datos\' + h1 + '.txt'

COPY TO &d1 TYPE sdf

CREATE TABLE R ( linea c(85))

APPEND FROM plspm_ref_w TYPE sdf

GOTO 2
h2= 'datos\\' + h1
instrucc = R.linea
instrucc=STRTRAN(instrucc,'ç',h2)
replace R.linea WITH instrucc
```

```
GOTO 7
h3= 'resR\\'+h1
instrucc2 = R.linea
instrucc2=STRTRAN(instrucc2,'c',h3)
replace R.linea WITH instrucc2

l1 ='c:\automatizac\ctrR\'+h1+'.r'

COPY TO &l1 TYPE sdf

a='"C:\Archivos de programa\R\R-2.14.0\bin\i386\Rscript.exe"'
r1 ='c:\automatizac\resR\'+h1+'.txt'

RUN/N &a &l1 &r1

WAIT TIMEOUT 0.01
?i
ENDFOR
```



---

**Anexo H.** Descripción de la sintaxis Visual FoxPro, para la extracción automatizada de información desde archivos de resultados de LISREL, al analizar el modelo con variables manifiestas reflectivas.

1.- Se determina el punto (".") como símbolo decimal (considerando que LISREL utiliza el punto para separar los enteros de los decimales).

```
SET POINT TO '.'
```

2.- Se define la cantidad de decimales a utilizar.

```
SET DECIMALS TO 6
```

3.- Se limpia la pantalla de salida de VFP y se cierran las tablas y bases de datos que se encuentren abiertas.

```
CLEAR  
CLOSE tables
```

4.- Se desactiva la solicitud de confirmación ante la modificación de archivos.

```
SET SAFETY OFF
```

5.- Se define la ruta del directorio de trabajo.

```
SET PATH TO c:\automatizac\reslisrel
```

6.- Se crea una tabla con los nombres de las variables que se extraerán de los archivos de resultados.

```
CREATE TABLE result1 (muestra i(4), LV1 n(8,2), LV2 f(8,2), LV3 f(8,2), ;  
LV4 f(8,2), LV5 f(8,2), LV6 f(8,2), LV7 f(8,2), ;  
LV8 f(8,2), LV9 f(8,2), BETA f(8,2), GAMMA f(8,2), ;  
RMSEA f(8,6), CFI f(8,3), SRMR f(8,3);  
)
```

7.- Se inicia un ciclo "for", determinándose la cantidad de ciclos a realizar.

```
FOR i = 1 TO 500
```

8.- Se crea la variable "h1", la cual tomará los valores de "i" definidos en el paso 7.

```
h1 = LTRIM(STR(i))
```

9.- Se agregan los nombres de los archivos de texto con los resultados LISREL a la variable "h1".

```
h1= 'Hoja'+h1+'.out'
```

10.- Se crea una tabla Excel con 7 columnas para recibir los resultados de LISREL.

```
CREATE TABLE excel (a c(14), b c(6), c c(5), d c(6), e c(27), f c(20), g  
c(7))
```

11.- Se añaden los contenidos de los archivos de resultados de LISREL a la tabla creada en el paso 10.

```
APPEND FROM &h1 TYPE SDF
```

- 12.- Se posiciona el control de ejecución a la línea 1 de la Tabla con los resultados de LISREL y se borran las 209 líneas siguientes.

```
GOTO 1
DELETE NEXT 209
pack
```

- 13.- Se posiciona el control de ejecución en la línea nº 1 y se selecciona el contenido de los 14 primeros caracteres de la columna "a".

```
GOTO 1
A = RIGHT(excel.a,14)
```

14. Se inicia un comando "if" para comprobar si existe una advertencia (asociada a PHI definida no positiva) en la línea nº 1. De existir dicha advertencia se ejecutarán los comandos que continúan al comando "then", que seleccionan el contenido de las variables RMSEA, CFI, SRMR, LV1, LV2, LV3, LV4, LV5, LV6, LV7, LV8, LV9, BETA y GAMMA.

```
IF A = ' W_A_R_N_I_N_G' THEN
GOTO 78
_RMSEA = VAL( right(excel.f,14) )
GOTO 98
_CFI = VAL( right(excel.e,5) )
GOTO 103
_SRMR = VAL( right(excel.e,10) )
GOTO 116
_LV1 = VAL(excel.b)
GOTO 117
_LV2 = VAL(excel.b)
GOTO 118
_LV3 = VAL(excel.b)
GOTO 119
_LV4 = VAL(excel.d)
GOTO 120
_LV5 = VAL(excel.d)
GOTO 121
_LV6 = VAL(excel.d)
GOTO 127
_LV7 = VAL(excel.b)
GOTO 128
_LV8 = VAL(excel.b)
GOTO 129
_LV9 = VAL(excel.b)
GOTO 136
_BETA = VAL(excel.b)
GOTO 142
_GAMMA = VAL(excel.b)
```

15. En el caso de no existir la advertencia en la línea nº 1 se ejecutarán los comandos ubicados luego del comando "else". Que posiciona el control de ejecución en la línea 33, para comprobar si existe una advertencia (asociada a PSI definida no positiva). De existir dicha advertencia se ejecutarán los comandos que continúan al comando "then", que seleccionan el contenido de las variables RMSEA, CFI, SRMR, LV1, LV2, LV3, LV4, LV5, LV6, LV7, LV8, LV9, BETA y GAMMA.

```
ELSE
GOTO 33
A = RIGHT(excel.a,14)
IF A = ' W_A_R_N_I_N_G' THEN
GOTO 78
_RMSEA = VAL( right(excel.f,14) )
```

```

GOTO 98
_CFI = VAL( right(excel.e,5) )
GOTO 103
_SRMR = VAL( right(excel.e,10) )
GOTO 116
_LV1 = VAL(excel.b)
GOTO 117
_LV2 = VAL(excel.b)
GOTO 118
_LV3 = VAL(excel.b)
GOTO 119
_LV4 = VAL(excel.d)
GOTO 120
_LV5 = VAL(excel.d)
GOTO 121
_LV6 = VAL(excel.d)
GOTO 127
_LV7 = VAL(excel.b)
GOTO 128
_LV8 = VAL(excel.b)
GOTO 129
_LV9 = VAL(excel.b)
GOTO 136
_BETA = VAL(excel.b)
GOTO 142
_GAMMA = VAL(excel.b)

```

- 16.- En el caso de no existir la advertencia en la línea nº 33 se ejecutarán los comandos ubicados entre el comando “else” y el comando “endif”, que seleccionan el contenido de las variables RMSEA, CFI, SRMR, LV1, LV2, LV3, LV4, LV5, LV6, LV7, LV8, LV9, BETA y GAMMA.

```

ELSE
GOTO 76
_RMSEA = VAL( right(excel.f,14) )
GOTO 96
_CFI = VAL( right(excel.e,5) )
GOTO 101
_SRMR = VAL( right(excel.e,10) )
GOTO 114
_LV1 = VAL(excel.b)
GOTO 115
_LV2 = VAL(excel.b)
GOTO 116
_LV3 = VAL(excel.b)
GOTO 117
_LV4 = VAL(excel.d)
GOTO 118
_LV5 = VAL(excel.d)
GOTO 119
_LV6 = VAL(excel.d)
GOTO 125
_LV7 = VAL(excel.b)
GOTO 126
_LV8 = VAL(excel.b)
GOTO 127
_LV9 = VAL(excel.b)
GOTO 134
_BETA = VAL(excel.b)
GOTO 140
_GAMMA = VAL(excel.b)
ENDIF

```

- 17.- Se finaliza el comando “if” iniciado en el paso 14.

```
ENDIF
```

18 - Se cierran las tablas activas de VFP.

```
CLOSE TABLES
```

19.- Se activa la Tabla que recibirá los resultados.

```
USE result1.dbf
```

20.- Se agrega un registro en blanco al final de la Tabla activa, para poder agregar la información relevante seleccionada a la Tabla activa.

```
APPEND BLANK
```

21.- Se reemplaza el contenido del campo (columna) "muestra", por el valor de la variable "i".

```
replace result1.muestra WITH i
```

22.- Se reemplaza el contenido de los campos correspondientes de la Tabla activa, por el valor de cada variable.

```
replace result1.lv1 WITH _LV1
replace result1.lv2 WITH _LV2
replace result1.lv3 WITH _LV3
replace result1.lv4 WITH _LV4
replace result1.lv5 WITH _LV5
replace result1.lv6 WITH _LV6
replace result1.lv7 WITH _LV7
replace result1.lv8 WITH _LV8
replace result1.lv9 WITH _LV9
replace result1.beta WITH _BETA
replace result1.gamma WITH _GAMMA
replace result1.RMSEA WITH _RMSEA
replace result1.CFI WITH _CFI
replace result1.SRMR WITH _SRMR
```

23.- Se finaliza el ciclo "for" iniciado en el paso 7.

```
ENDFOR
```

24.- Se muestra en pantalla la última Tabla activa (result1.dbf)

```
BROWSE
```

25.- Se copia el contenido de la Tabla activa a un archivo Excel.

```
COPY TO res_lisrel_500s_a0_policorica_uls_well.xls TYPE XL5
```

Sintaxis Visual FoxPro:

```
SET POINT TO '.'
SET DECIMALS TO 6
CLEAR
CLOSE tables
SET SAFETY OFF
SET PATH TO c:\automatizac\reslisrel

CREATE TABLE result1 (muestra i(4), LV1 n(8,2), LV2 f(8,2), LV3 f(8,2), ;
LV4 f(8,2), LV5 f(8,2), LV6 f(8,2), LV7 f(8,2), ;
LV8 f(8,2), LV9 f(8,2), BETA f(8,2), GAMMA f(8,2), ;
```

```

RMSEA f(8,6), CFI f(8,3), SRMR f(8,3);
)

FOR i = 1 TO 500
h1 = LTRIM(STR(i))
h1 = 'Hoja'+h1+'.out'

CREATE TABLE excel (a c(14), b c(6), c c(5), d c(6), e c(27), f c(20), g c(7))
APPEND FROM &h1 TYPE SDF

GOTO 1
DELETE NEXT 209
pack

GOTO 1
A = RIGHT(excel.a,14)
IF A = ' W_A_R_N_I_N_G' THEN
GOTO 78
_RMSEA = VAL( right(excel.f,14) )
GOTO 98
_CFI = VAL( right(excel.e,5) )
GOTO 103
_SRMR = VAL( right(excel.e,10) )
GOTO 116
_LV1 = VAL(excel.b)
GOTO 117
_LV2 = VAL(excel.b)
GOTO 118
_LV3 = VAL(excel.b)
GOTO 119
_LV4 = VAL(excel.d)
GOTO 120
_LV5 = VAL(excel.d)
GOTO 121
_LV6 = VAL(excel.d)
GOTO 127
_LV7 = VAL(excel.b)
GOTO 128
_LV8 = VAL(excel.b)
GOTO 129
_LV9 = VAL(excel.b)
GOTO 136
_BETA = VAL(excel.b)
GOTO 142
_GAMMA = VAL(excel.b)
ELSE
GOTO 33
A = RIGHT(excel.a,14)
IF A = ' W_A_R_N_I_N_G' THEN
GOTO 78
_RMSEA = VAL( right(excel.f,14) )
GOTO 98
_CFI = VAL( right(excel.e,5) )
GOTO 103
_SRMR = VAL( right(excel.e,10) )
GOTO 116
_LV1 = VAL(excel.b)
GOTO 117
_LV2 = VAL(excel.b)
GOTO 118
_LV3 = VAL(excel.b)
GOTO 119
_LV4 = VAL(excel.d)
GOTO 120
_LV5 = VAL(excel.d)
GOTO 121
_LV6 = VAL(excel.d)
GOTO 127

```

```

    _LV7 = VAL(excel.b)
    GOTO 128
    _LV8 = VAL(excel.b)
    GOTO 129
    _LV9 = VAL(excel.b)
    GOTO 136
    _BETA = VAL(excel.b)
    GOTO 142
    _GAMMA = VAL(excel.b)
ENDIF
ELSE
    GOTO 76
    _RMSEA = VAL( right(excel.f,14) )
    GOTO 96
    _CFI = VAL( right(excel.e,5) )
    GOTO 101
    _SRMR = VAL( right(excel.e,10) )
    GOTO 114
    _LV1 = VAL(excel.b)
    GOTO 115
    _LV2 = VAL(excel.b)
    GOTO 116
    _LV3 = VAL(excel.b)
    GOTO 117
    _LV4 = VAL(excel.d)
    GOTO 118
    _LV5 = VAL(excel.d)
    GOTO 119
    _LV6 = VAL(excel.d)
    GOTO 125
    _LV7 = VAL(excel.b)
    GOTO 126
    _LV8 = VAL(excel.b)
    GOTO 127
    _LV9 = VAL(excel.b)
    GOTO 134
    _BETA = VAL(excel.b)
    GOTO 140
    _GAMMA = VAL(excel.b)
ENDIF
ENDIF

CLOSE TABLES

USE result1.dbf
APPEND BLANK

replace result1.muestra WITH i
replace result1.LV1 WITH _LV1
replace result1.lv2 WITH _LV2
replace result1.lv3 WITH _LV3
replace result1.lv4 WITH _LV4
replace result1.lv5 WITH _LV5
replace result1.lv6 WITH _LV6
replace result1.lv7 WITH _LV7
replace result1.lv8 WITH _LV8
replace result1.lv9 WITH _LV9
replace result1.beta WITH _BETA
replace result1.gamma WITH _GAMMA
replace result1.RMSEA WITH _RMSEA
replace result1.CFI WITH _CFI
replace result1.SRMR WITH _SRMR

ENDFOR

BROWSE

COPY TO res_lisrel_500s_a0_policorica_uls_well.xls TYPE XL5

```

---

**Anexo I.** Descripción de la sintaxis Visual FoxPro, para la extracción automatizada de información desde archivos de resultados de R, al analizar el modelo con variables manifiestas reflectivas.

- 1.- Se determina el punto (".") como símbolo decimal (considerando que R utiliza el punto para separar los enteros de los decimales).

```
SET POINT TO '.'
```

- 2.- Se define la cantidad de decimales a utilizar.

```
SET DECIMALS TO 6
```

- 3.- Se limpia la pantalla de salida de VFP y se cierran las tablas y bases de datos que se encuentren abiertas.

```
CLEAR  
CLOSE tables
```

- 4.- Se desactiva la solicitud de confirmación ante la modificación de archivos.

```
SET SAFETY OFF
```

- 5.- Se define la ruta del directorio de trabajo.

```
SET PATH TO C:\Automatizac\resR
```

- 6.- Se crea una tabla con los nombres de las variables que se extraerán de los archivos de resultados.

```
CREATE TABLE result1 (muestra i(4), LV1 n(8,2), LV2 f(8,2), LV3 f(8,2), ;  
LV4 f(8,2), LV5 f(8,2), LV6 f(8,2), LV7 f(8,2), ;  
LV8 f(8,2), LV9 f(8,2), BETA f(8,2), GAMMA f(8,2), ;  
GoF f(8,3);  
)
```

- 7.- Se inicia un ciclo "for", determinándose la cantidad de ciclos a realizar.

```
FOR i = 1 TO 500
```

- 8.- Se crea la variable "h1", la cual tomará los valores de "i" definidos en el punto 6.

```
h1 = LTRIM(STR(i))
```

- 9.- Se agregan los nombres de los archivos de texto con los resultados R a la variable "h1".

```
h1 = 'Hoja'+h1+'.txt'
```

- 10.- Se crea una tabla Excel con 7 columnas para recibir los resultados de R.

```
CREATE TABLE excel (a c(8), b c(10), c c(10), d c(10), e c(27), f c(20), g  
c(7))
```

- 11.- Se añaden los contenidos de los archivos de resultados de R a la tabla creada en el paso 10.

```
APPEND FROM &h1 TYPE SDF
```

- 12.- Se posiciona el control de ejecución a la línea 1 de la Tabla con los resultados de R y se borran las 35 líneas siguientes.

```
GOTO 1  
DELETE NEXT 35  
PACK
```

- 13.- Se posiciona el control de ejecución en la línea nº 1 de la Tabla de resultados y se asigna el valor numérico de la columna “c” de la Tabla, a la variable “\_LV7”.

```
GOTO 1  
_LV7 = VAL(excel.c)
```

- 14.- Se asignan los valores numéricos, de las columnas correspondientes de la Tabla de resultados, a cada variable.

```
GOTO 2  
_LV8 = VAL(excel.c)  
GOTO 3  
_LV9 = VAL(excel.c)  
GOTO 5  
_LV1 = VAL(excel.c)  
GOTO 6  
_LV2 = VAL(excel.c)  
GOTO 7  
_LV3 = VAL(excel.c)  
GOTO 9  
_LV4 = VAL(excel.c)  
GOTO 10  
_LV5 = VAL(excel.c)  
GOTO 11  
_LV6 = VAL(excel.c)  
GOTO 35  
_GAMMA = VAL(RIGHT(excel.b,5))  
GOTO 41  
_BETA = VAL(RIGHT(excel.b,5))  
GOTO 61  
_GoF = VAL(RIGHT(excel.b,4))
```

- 15 - Se cierran las tablas activas de VFP.

```
CLOSE TABLES
```

- 16.- Se activa la Tabla que recibirá los resultados.

```
USE result1.dbf
```

- 17.- Se agrega un registro en blanco al final de la Tabla activa, para poder agregar la información relevante seleccionada a la Tabla activa.

```
APPEND BLANK
```

- 18.- Se reemplaza el contenido del campo (columna) “muestra”, por el valor de la variable “i”.

```
replace result1.muestra WITH i
```

- 19.- Se reemplaza el contenido de los campos correspondientes de la Tabla activa, por el valor de cada variable.



---

```

replace result1.lv1 WITH _LV1
replace result1.lv2 WITH _LV2
replace result1.lv3 WITH _LV3
replace result1.lv4 WITH _LV4
replace result1.lv5 WITH _LV5
replace result1.lv6 WITH _LV6
replace result1.lv7 WITH _LV7
replace result1.lv8 WITH _LV8
replace result1.lv9 WITH _LV9
replace result1.beta WITH _BETA
replace result1.gamma WITH _GAMMA
replace result1.gof WITH _GoF

```

20.- Se finaliza el ciclo “for”.

```
ENDFOR
```

21.- Se muestra en pantalla la última Tabla activa (result1.dbf)

```
BROWSE
```

22.- Se copia el contenido de la Tabla activa a un archivo Excel.

```
COPY TO resultR_500s_a0.xls TYPE XL5
```

Sintaxis Visual FoxPro:

```

SET POINT TO '.'
SET DECIMALS TO 6
CLEAR
CLOSE tables
SET SAFETY OFF
SET PATH TO C:\Automatizac\resR

CREATE TABLE result1 (muestra i(4), LV1 n(8,2), LV2 f(8,2), LV3 f(8,2), ;
LV4 f(8,2), LV5 f(8,2), LV6 f(8,2), LV7 f(8,2), ;
LV8 f(8,2), LV9 f(8,2), BETA f(8,2), GAMMA f(8,2), ;
GoF f(8,2);
)

FOR i = 1 TO 500

h1 = LTRIM(STR(i))
h1 = 'Hoja'+h1+'.txt'

CREATE TABLE excel (a c(8), b c(10), c c(10), d c(9), e c(27), f c(20), g
c(7))

APPEND FROM &h1 TYPE SDF

GOTO 1
DELETE NEXT 35
pack

GOTO 1
_LV7 = VAL(excel.c)
GOTO 2
_LV8 = VAL(excel.c)
GOTO 3
_LV9 = VAL(excel.c)
GOTO 5
_LV1 = VAL(excel.c)
GOTO 6

```

## Anexos

---

```
_LV2 = VAL(excel.c)
GOTO 7
_LV3 = VAL(excel.c)
GOTO 9
_LV4 = VAL(excel.c)
GOTO 10
_LV5 = VAL(excel.c)
GOTO 11
_LV6 = VAL(excel.c)
GOTO 35
_GAMMA = VAL(RIGHT(excel.b,5))
GOTO 41
_BETA = VAL(RIGHT(excel.b,5))
GOTO 61
_GoF = VAL(RIGHT(excel.b,4))

CLOSE TABLES

USE result1.dbf
APPEND BLANK

replace result1.muestra WITH i
replace result1.lv1 WITH _LV1
replace result1.lv2 WITH _LV2
replace result1.lv3 WITH _LV3
replace result1.lv4 WITH _LV4
replace result1.lv5 WITH _LV5
replace result1.lv6 WITH _LV6
replace result1.lv7 WITH _LV7
replace result1.lv8 WITH _LV8
replace result1.lv9 WITH _LV9
replace result1.beta WITH _BETA
replace result1.gamma WITH _GAMMA
replace result1.gof WITH _GoF

ENDFOR

BROWSE

COPY TO result_R_500s_a0.xls TYPE XL5
```

---

**Anexo J.** Descripción de la sintaxis para la simulación de datos con Matlab, a partir de la generación de variables aleatorias del modelo con variables manifiestas formativas y reflectivas, modelo más especificado.

- 1.- Se borran registros del espacio de trabajo, se define un valor de semilla para asegurar que los datos generados sean siempre los mismos.

```
clear
semilla=300000;
randn('state',semilla);
```

- 2.- Se definen valores para: el número de muestras, número de sujetos, número de factores, número de ítems por factor, y el número total de ítems.

```
nmuestras=100; nsuj=500; nfact=3; nitemsfact=3; Nitems=nfact*nitemsfact;
```

- 3.- Se definen los pesos que tendrán las variables manifiestas de cada factor.

```
lambdasf1=0.4; lambdasf2=0.7; lambdasf3=0.7;
```

- 4.- Se definen los valores de Beta, coeficientes de regresión de variables endógenas sobre variables endógena a las que afectan.

```
beta21=0.6; beta32=0.6;
```

- 5.- Se definen medias para la posterior generación de variables epsilon, Zeta y x.

```
mediase=0; mediasz=0;
mediasx=[0 0 0];
```

- 6.- Se define la matriz Phi, con las varianzas y covarianzas entre las variables manifiestas asociadas a las variables latentes exógenas.

```
sigmax=[1 0.4 0.4; 0.4 1 0.4; 0.4 0.4 1];
```

- 7.- Se generan los vectores columna que recibirán los coeficientes de regresión de entre las variables latentes Eta1 sobre Eta2 (rE1E2) y Eta2 sobre Eta3 (rE2E3), para comprobar los valores de Beta.

```
rE1E2=zeros(nmuestras,1);
rE2E3=zeros(nmuestras,1);
```

- 8.- Se construye una matriz tridimensional (D3) para guardar los datos generados, con x=cantidad de sujetos, y=cantidad de ítems; y z=cantidad de muestras.

```
D3=zeros(nsuj,Nitems,nmuestras);
```

- 9.- Se inicia un ciclo "for" en el cual se determina que la cantidad de ciclos será igual a la cantidad de muestras definida en el paso 2.

```
for i=1:nmuestras;
```

- 10.- Se generan los valores de las variables manifiestas x.

```
x=mvnrnd(mediasx,sigmax,nsuj);
```

11.- Se generan los valores de la varianza de Zeta1 ( $\sigma_{z1}$ ), para que Eta1 tenga varianza igual a 1. Además se generan los varlores de Zeta1 y Eta1.

```
sigmaz1=(1-((3*lambdaf1^2)+3*(2*(sigma(1,2)^3))));
Zeta1=mvnrnd(mediasz,sigmaz1,nsuj);
Eta1=(lambdaf1*x(:,1))+(lambdaf1*x(:,2))+(lambdaf1*x(:,3))+Zeta1;
```

12.- Se generan los valores de la varianza de Zeta2 ( $\sigma_{z2}$ ), para que Eta2 tenga varianza igual a 1. Además se generan los varlores de Zeta2.

```
sigmaz2=(1-(beta21^2));
Zeta2=mvnrnd(mediasz,sigmaz2,nsuj);
```

13.- Se calcula la varianza de Theta-Epsilon ( $\sigma_{\epsilon 1}$ ), necesaria para que las variables manifiestas asociadas a Eta2 tengan varianza igual a 1. Se generan los valores de los errores asociados a las variables manifiestas asociadas a Eta2.

```
sigmae1=(1-
((3*(lambdaf2^2*beta21^2*lambdaf1^2))+(lambdaf2^2*beta21^2*var(Zeta1))+(l
ambdaf2^2*var(Zeta2))));
epsilon1=mvnrnd(mediase,sigmae1,nsuj);
epsilon2=mvnrnd(mediase,sigmae1,nsuj);
epsilon3=mvnrnd(mediase,sigmae1,nsuj);
```

14.- Se generan los valores de las variables manifiestas “y” asociadas a Eta2.

```
y1=(lambdaf2*beta21*lambdaf1*x(:,1))+(lambdaf2*beta21*lambdaf1*x(:,2))+(
lambdaf2*beta21*lambdaf1*x(:,3))+(lambdaf2*beta21*Zeta1)+(lambdaf2*Zeta2
)+(epsilon1);

y2=(lambdaf2*beta21*lambdaf1*x(:,1))+(lambdaf2*beta21*lambdaf1*x(:,2))+(
lambdaf2*beta21*lambdaf1*x(:,3))+(lambdaf2*beta21*Zeta1)+(lambdaf2*Zeta2
)+(epsilon2);

y3=(lambdaf2*beta21*lambdaf1*x(:,1))+(lambdaf2*beta21*lambdaf1*x(:,2))+(
lambdaf2*beta21*lambdaf1*x(:,3))+(lambdaf2*beta21*Zeta1)+(lambdaf2*Zeta2
)+(epsilon3);
```

15. Se generan los valores de la varianza de Zeta3 ( $\sigma_{z3}$ ), para que Eta3 tenga varianza igual a 1. Además se generan los varlores de Zeta3.

```
sigmaz3=(1-((beta32^2)));
Zeta3=mvnrnd(mediasz,sigmaz3,nsuj);
```

16.- Se calcula la varianza de Theta-Epsilon ( $\sigma_{\epsilon 2}$ ), necesaria para que las variables manifiestas asociadas a Eta3 tengan varianza igual a 1. Se generan los valores de los errores asociados a las variables manifiestas asociadas a Eta3.

```
sigmae2=(1-
((3*(lambdaf3^2*beta32^2*beta21^2*lambdaf1^2))+(lambdaf3^2*beta32^2*beta2
1^2*var(Zeta1))+(lambdaf3^2*beta32^2*var(Zeta2))+(lambdaf3^2*var(Zeta3))));
epsilon4=mvnrnd(mediase,sigmae2,nsuj);
epsilon5=mvnrnd(mediase,sigmae2,nsuj);
epsilon6=mvnrnd(mediase,sigmae2,nsuj);
```

17.- Se calculan los valores de las variables latentes Eta 2 y Eta3.

```
Eta2=(beta21*Eta1)+(Zeta2);
Eta3=(beta32*beta21*lambdaf1*x(:,1))+(beta32*beta21*lambdaf1*x(:,2))+(beta
32*beta21*lambdaf1*x(:,3))+(beta32*beta21*Zeta1)+(beta32*Zeta2)+(Zeta3);
```

18.- Se generan las variables manifiestas “y” asociadas a la variable latente Eta3.

```
y4=(lambdasf3*beta32*beta21*lambdasf1*x(:,1))+(lambdasf3*beta32*beta21*lambdasf1*x(:,2))+(lambdasf3*beta32*beta21*lambdasf1*x(:,3))+(lambdasf3*beta32*beta21*Zeta1)+(lambdasf3*beta32*Zeta2)+(lambdasf3*Zeta3)+(epsilon4);
```

```
y5=(lambdasf3*beta32*beta21*lambdasf1*x(:,1))+(lambdasf3*beta32*beta32*lambdasf1*x(:,2))+(lambdasf3*beta32*beta21*lambdasf1*x(:,3))+(lambdasf3*beta32*beta21*Zeta1)+(lambdasf3*beta32*Zeta2)+(lambdasf3*Zeta3)+(epsilon5);
```

```
y6=(lambdasf3*beta32*beta21*lambdasf1*x(:,1))+(lambdasf3*beta32*beta32*lambdasf1*x(:,2))+(lambdasf3*beta32*beta21*lambdasf1*x(:,3))+(lambdasf3*beta32*beta21*Zeta1)+(lambdasf3*beta32*Zeta2)+(lambdasf3*Zeta3)+(epsilon6);
```

19.- Se calculan las regresiones de Eta1 sobre Eta2 y de Eta2 sobre Eta3.

```
rE1E2(i)=regress(Eta2,Eta1);
rE2E3(i)=regress(Eta3,Eta2);
```

20.- Se crea la matriz de datos (datos) con las variables manifiestas.

```
datos=horzcat(y1,y2,y3,y4,y5,y6,x(:,1),x(:,2),x(:,3));
```

21.- Se inicia un ciclo “for” en el cual se categorizan los datos continuos en 5 puntos, de acuerdo a los umbrales señalados por Garrido, Abad, Ponsoda (2011).

```
for j=1:Nitems
    for h=1:nsuj
        if datos(h,j)<=0.68
            datos(h,j)=1;
        elseif 0.68<datos(h,j) && datos(h,j)<=1.00
            datos(h,j)=2;
        elseif 1.00<datos(h,j) && datos(h,j)<=1.34
            datos(h,j)=3;
        elseif 1.34<datos(h,j) && datos(h,j)<=1.77
            datos(h,j)=4;
        else
            datos(h,j)=5;
        end
    end
end
```

22.- Se guardan los datos en la matriz tridimensional D3.

```
D3(:,:,i)=datos;
```

23.- Se exportan los datos a una planilla Excel, con el nombre “datos\_500s\_a2\_form\_ord\_F1” en la cual, cada hoja corresponde a una muestra de datos ordinales de 5 puntos y se finaliza el ciclo “for” iniciado en el punto 9.

```
xlswrite('datos_500s_a2_form_ord_F1.xls',D3(:,:,i),i)
end
```

Sintaxis Matlab:

```
clear
semilla=300000;
randn('state',semilla);
nmuestras=100; nsuj=500; itemsfact=3; fact=3; Nitems=itemsfact*fact;
lambdasf1=0.4; lambdasf2=0.7; lambdasf3=0.7;
beta21=0.6; beta32=0.6;
```

---

```

mediase=0; mediasz=0; mediasx=[0 0 0];
sigmax=[1 0.4 0.4; 0.4 1 0.4; 0.4 0.4 1];
rE1E2=zeros(nmuestras,1);
rE2E3=zeros(nmuestras,1);
D3=zeros(nsuj,Nitems,nmuestras);
for i=1:nmuestras;
    x=mvnrnd(mediasx,sigmax,nsuj);
    sigmaz1=(1-(3*lambdaf1^2)+3*(2*(sigmax(1,2)^2)));
    Zeta1=mvnrnd(mediasz,sigmaz1,nsuj);
    Eta1=(lambdaf1*x(:,1)+(lambdaf1*x(:,2))+(lambdaf1*x(:,3))+Zeta1);
    sigmaz2=(1-(beta21^2));
    Zeta2=mvnrnd(mediasz,sigmaz2,nsuj);
    sigmae1=(1-
    ((3*(lambdaf2^2*beta21^2*lambdaf1^2)+(lambdaf2^2*beta21^2*var(Zeta1)+(lambdaf2^2*var(Zeta2)))));
    epsilon1=mvnrnd(mediase,sigmae1,nsuj);
    epsilon2=mvnrnd(mediase,sigmae1,nsuj);
    epsilon3=mvnrnd(mediase,sigmae1,nsuj);
    y1=(lambdaf2*beta21*lambdaf1*x(:,1)+(lambdaf2*beta21*lambdaf1*x(:,2))+(lambdaf2*beta21*lambdaf1*x(:,3))+(lambdaf2*beta21*Zeta1)+(lambdaf2*beta21*Zeta2)+(epsilon1);
    y2=(lambdaf2*beta21*lambdaf1*x(:,1)+(lambdaf2*beta21*lambdaf1*x(:,2))+(lambdaf2*beta21*lambdaf1*x(:,3))+(lambdaf2*beta21*Zeta1)+(lambdaf2*beta21*Zeta2)+(epsilon2);
    y3=(lambdaf2*beta21*lambdaf1*x(:,1)+(lambdaf2*beta21*lambdaf1*x(:,2))+(lambdaf2*beta21*lambdaf1*x(:,3))+(lambdaf2*beta21*Zeta1)+(lambdaf2*beta21*Zeta2)+(epsilon3);
    sigmaz3=(1-(beta32^2));
    Zeta3=mvnrnd(mediasz,sigmaz3,nsuj);
    sigmae2=(1-
    ((3*(lambdaf3^2*beta32^2*beta21^2*lambdaf1^2)+(lambdaf3^2*beta32^2*beta21^2*var(Zeta1)+(lambdaf3^2*beta32^2*var(Zeta2)))+(lambdaf3^2*var(Zeta3)))));
    epsilon4=mvnrnd(mediase,sigmae2,nsuj);
    epsilon5=mvnrnd(mediase,sigmae2,nsuj);
    epsilon6=mvnrnd(mediase,sigmae2,nsuj);
    Eta2=(beta21*Eta1)+(Zeta2);

Eta3=(beta32*beta21*lambdaf1*x(:,1)+(beta32*beta21*lambdaf1*x(:,2))+(beta32*beta21*lambdaf1*x(:,3))+(beta32*beta21*Zeta1)+(beta32*beta21*Zeta2)+(Zeta3);

y4=(lambdaf3*beta32*beta21*lambdaf1*x(:,1)+(lambdaf3*beta32*beta21*lambdaf1*x(:,2))+(lambdaf3*beta32*beta21*lambdaf1*x(:,3))+(lambdaf3*beta32*beta21*Zeta1)+(lambdaf3*beta32*Zeta2)+(lambdaf3*beta32*Zeta3)+(epsilon4);

y5=(lambdaf3*beta32*beta21*lambdaf1*x(:,1)+(lambdaf3*beta32*beta21*lambdaf1*x(:,2))+(lambdaf3*beta32*beta21*lambdaf1*x(:,3))+(lambdaf3*beta32*beta21*Zeta1)+(lambdaf3*beta32*Zeta2)+(lambdaf3*beta32*Zeta3)+(epsilon5);

y6=(lambdaf3*beta32*beta21*lambdaf1*x(:,1)+(lambdaf3*beta32*beta21*lambdaf1*x(:,2))+(lambdaf3*beta32*beta21*lambdaf1*x(:,3))+(lambdaf3*beta32*beta21*Zeta1)+(lambdaf3*beta32*Zeta2)+(lambdaf3*beta32*Zeta3)+(epsilon6);
    rE1E2(i)=regress(Eta2,Eta1);
    rE2E3(i)=regress(Eta3,Eta2);
    datos=horzcat(y1,y2,y3,y4,y5,y6,x(:,1),x(:,2),x(:,3));
    for j=1:Nitems
        for h=1:nsuj
            if datos(h,j)<=0.68
                datos(h,j)=1;
            elseif 0.68<datos(h,j) && datos(h,j)<=1.00
                datos(h,j)=2;
            elseif 1.00<datos(h,j) && datos(h,j)<=1.34
                datos(h,j)=3;
            elseif 1.34<datos(h,j) && datos(h,j)<=1.77
                datos(h,j)=4;
            else
                datos(h,j)=5;
            end
        end
    end
end

```

---

```
end
D3(:,:,i)=datos;
xlswrite('datos_50s_a2_form_ord_F1.xls',D3(:,:,i),i)
end
```

**Anexo K.** Descripción de la sintaxis LISREL para el análisis ACOV del modelo con variables manifiestas formativas y reflectivas, más especificado.

1.- Se establece el título (TI) de la sintaxis.

```
TI Ref_Form
```

2.- Se define los datos (DA), en cuanto al número de variables manifiestas del modelo (NI), el tamaño de la muestra (NO) y que la matriz analizada (MA) será la de correlaciones de Pearson (KM).

```
DA NI=9 NO=500 MA=KM
```

3.- Se determina la ruta (FI) para acceder la matriz de datos brutos (RA).

```
RA='C:\Automatizac\ç.txt'
```

4.- Se determina el orden en que serán ingresadas las variables al modelo (SE), comenzando por las variables manifiestas asociadas a las variables latentes endógenas.

```
SE  
1 2 3 4 5 6 7 8 9 /
```

5.- Se especifica el modelo (MO), determinando la cantidad de variables manifiestas exógenas (NX), la cantidad de variables manifiestas endógenas (NY), la cantidad de variables latentes exógenas (NK) y la cantidad de variables latentes endógenas (NE). Luego se especifica que Beta (BE) y Gamma son completas (FU) y libres (FR), la matriz Phi (PH) es simétrica (SY) libre (FR) y que las matrices Psi (PS), Theta Epsilon (TE) y Theta Delta (TD) son diagonales (DI) y libres (FR).

```
MO NX=3 NY=6 NK=3 NE=3 BE=FU,FI GA=FU,FI PH=SY,FR PS=DI,FR TE=DI,FR  
TD=DI,FR
```

6.- Se definen los nombres de las variables latentes endógenas (LE) y las variables latentes exógenas (LK).

```
LE  
'VL1' 'VL 2' 'VL 3'  
LK  
FORM7 FORM8 FORM9
```

7.- Se fija a 0 (FI) los valores de la varianza de Zeta1 (PS(1,1)) y las varianzas de la matriz Theta-Delta.

```
FI PS(1,1) TD(1,1) TD(2,2) TD(3,3)
```

8.- Se definen los parámetros libres (FR) de Lambda que serán estimados.

```
FR LY(2,2) LY(3,2) LY(5,3) LY(6,3)
```

9.- Se definen los parámetros libres de Beta.

```
FR BE(2,1) BE(3,2)
```

10.- Se fijan los parámetros de Gamma que no serán estimados.

```
FI GA(2,1) GA(2,2) GA(2,3) GA(3,1) GA(3,2) GA(3,3)
```



11.- Se fijan a 1 los valores que dan métrica a los parámetros que darán métrica a las variables latentes.

```
VA 1 LY(1,2)
VA 1 LY(4,3)
VA 1 LX(1,1)
VA 1 LX(2,2)
VA 1 LX(3,3)
VA 1 GA(1,1)
```

12.- Opcionalmente, se puede definir la representación del diagrama estructural.

```
PD
```

13.- Se define que para la obtención de resultados (OU) el método de estimación de parámetros (ME) será Máxima Verosimilitud (ML), se solicita la solución estandarizada (SS) y se define el número máximo de iteraciones (AD) para comprobar la admisibilidad de la solución.

```
OU ME=ML SS AD=100
```

**Sintaxis LISREL:**

```
TI Ref_Form
DA NI=9 NO=500 MA=KM
RA='C:\Automatizac\ç.txt'
SE
1 2 3 4 5 6 7 8 9 /
MO NX=3 NY=6 NK=3 NE=3 BE=FU,FI GA=FU,FI PH=SY,FR PS=DI,FR TE=DI,FR TD=DI,FR
LE
'VL1' 'VL 2' 'VL 3'
LK
FORM7 FORM8 FORM9
FI PS(1,1) TD(1,1) TD(2,2) TD(3,3)
FR LY(2,2) LY(3,2) LY(5,3) LY(6,3)
FR BE(2,1) BE(3,2)
FI GA(2,1) GA(2,2) GA(2,3) GA(3,1) GA(3,2) GA(3,3)
VA 1 LY(1,2)
VA 1 LY(4,3)
VA 1 LX(1,1)
VA 1 LX(2,2)
VA 1 LX(3,3)
VA 1 GA(1,1)
PD
OU ME=ML SS AD=100
```

**Anexo L.** Descripción de la sintaxis R para el análisis PLS-PM del modelo con variables manifiestas formativas y reflectivas, más especificado.

1.- Se abre el paquete “plspm”.

```
library(plspm)
```

2.- Se asignan los datos al objeto denominado “A”.

```
A <-read.table("C:\\Automatizac\\ç.txt")
```

3.- Se especifica la matriz Beta y se asigna al objeto denominado “inner”.

```
inner <- matrix(c(0,0,0,1,0,0,0,1,0),3,3,byrow=TRUE)
```

4.- Se definen la matriz Lambda y se asocia al objeto “outer”, comenzando por las variables manifiestas asociadas a la variable latente exógena.

```
outer <- list(c(7,8,9),c(1,2,3),c(4,5,6))
```

5.- Se define la modalidad formativa (“B”) para el primer bloque de variables manifiestas y la modalidad reflejiva (“A”) para los otros dos bloques de variables manifiestas y se asignan la objeto denominado “mod”.

```
mod <- c("B","A","A")
```

6.- Se realiza el análisis PLS-PM utilizando los objetos definidos en los pasos 2, 3, 4 y 5, determinando el esquema Path weighting para la estimación de parámetros; y se asigna al objeto denominado “C”.

```
C <-plspm(A,inner=inner,outer=outer,modes=mod,scheme="path")
```

7.- Se define la ruta del archivo en que serán guardados los resultados.

```
sink(file="C:\\Automatizac\\ç.txt")
```

8.- Se solicita el resumen del análisis PLS-PM del objeto “C”

```
summary(C)
```

9.- Se cierra la función que guarda los archivos especificados en el paso anterior.

```
sink()
```

**Sintaxis R:**

```
library(plspm)
A<-read.table("C:\\Automatizac\\ç.txt")
inner <- matrix(c(0,0,0,1,0,0,0,1,0),3,3,byrow=TRUE)
outer <- list(c(7,8,9),c(1,2,3),c(4,5,6))
mod <- c("B","A","A")
C<-plspm(A,inner=inner,outer=outer,modes=mod,scheme="path")
sink(file="C:\\Automatizac\\ç.txt")
summary(C)
sink()
```

---

**Anexo M.** Descripción de la sintaxis Visual FoxPro, para la ejecución automatizada de análisis ACOV con LISREL, modelo con variables manifiestas formativas y reflectivas, más especificado.

- 1.- Se borra la ventana de salida y se cierran archivos abiertos.

```
CLEAR
CLOSE ALL
```

- 2.- Se desactiva la solicitud de confirmación para modificar archivos.

```
SET SAFETY OFF
```

- 3.- Se establece la ruta del directorio de trabajo.

```
SET PATH TO c:\automatizac
```

- 4.- Se inicia un ciclo “for”, determinándose la cantidad de ciclos a realizar.

```
FOR i = 1 TO 100
```

- 5.- Se crea la variable “h1”, la cual tomará los valores de “i” definidos en el punto 4.

```
h1 = LTRIM(STR(i))
```

- 6.- Se agrega la palabra “Hoja” a los valores que tomará la variable “h1”.

```
h1= 'Hoja'+h1
```

- 7.- Se crea una tabla denominada “excel” con 9 columnas, cada una con 1 carácter.

```
CREATE TABLE excel (a i(1), b i(1), c i(1), d i(1), e i(1), f i(1), g i(1),
h i(1), i i(1) )
```

- 8.- Se especifica el nombre del archivo del cual se extraen los datos, los cuales se añaden a la tabla “excel” creada en el paso anterior.

```
APPEND FROM datos_50s_a2_form_ord_F1.xls TYPE xl8 SHEET &h1
```

- 9.- Se crea la variable “d1”, que recibe el nombre de los archivos de datos con formato de texto.

```
d1 = 'c:\automatizac\datos\' + h1 + '.txt'
```

- 10.- Se copian los datos en archivos de texto creados en paso 9.

```
COPY TO &d1 TYPE sdf
```

- 11.- Se crea una tabla denominada “lisrel” con 80 caracteres.

```
CREATE TABLE lisrel ( linea c(80))
```

- 12.- Se especifica el nombre del archivo del cual se extrae la sintaxis LISREL, la que se añade a la tabla “lisrel” creada en el paso anterior.

```
APPEND FROM formMLw.LS8 TYPE sdf
```

- 13.- Se posiciona el control de ejecución en la línea nº 3 de la Tabla con la sintaxis de LISREL (en la que se especifica la ruta para acceder a los datos, con el nombre de archivo definido por “ç”).

```
GOTO 3
```

- 14.- Se crea la variable “h2”, que contiene parte de la ruta para acceder al archivo que contiene los datos.

```
h2= 'datos\'+h1
```

- 15.- Se crea la variable “instrucc”, que constituye una línea de la tabla “lisrel”.

```
instrucc2 = lisrel.linea
```

- 16.- Se cambia el nombre del archivo que contiene los datos (“ç”) por los valores de la variable “h2”.

```
instrucc2=strtran(instrucc, 'ç', h2)
```

- 17.- Se reemplaza el contenido de la línea nº 3 del archivo de sintaxis LISREL por el contenido de “instrucc” (con el nombre de los archivos de datos).

```
replace lisrel.linea WITH instrucc
```

- 18.- Se crea la variable “l1”, que recibe la ruta con el nombre del archivo LISREL que tiene la línea nº 3 reemplazada.

```
l1 ='c:\automatizac\ctrlisrel\'+h1+'.lpj'
```

- 19.- Se graba el archivo LISREL con el nombre archivo definido en el paso 18.

```
COPY TO &l1 TYPE sdf
```

- 20.- Se crea la variable “a”, que recibe la ruta con el nombre del programa batch de LISREL.

```
a="'C:\Archivos de programa\LISREL88\lisrel88.EXE"'
```

- 21.- Se crea la variable “r1”, que recibe la ruta con el nombre del archivo para guardar resultados.

```
r1 ='c:\automatizac\reslisrel\'+h1+'.out'
```

- 22.- Se ejecuta LISREL con la sintaxis definida por “l1” y guardando los resultados en “r1”.

```
RUN/N&a &l1 &r1
```

- 23.- Se define el tiempo de pausa para continuar con el siguiente ciclo.

```
WAIT TIMEOUT 0.001
```

- 24.- Se muestra en pantalla el valor de “i”, correspondiente al ciclo “for” que se está ejecutando.

```
?i
```

---

25.- Se finaliza el ciclo "for".

```
ENDFOR
```

Sintaxis Visual FoxPro:

```
CLEAR
CLOSE ALL
SET SAFETY OFF
SET PATH TO c:\automatizac

FOR i = 1 TO 100

h1 = LTRIM(STR(i))
h1= 'Hoja'+h1

CREATE TABLE excel (a i(1), b i(1), c i(1), d i(1), e i(1), f i(1), g i(1), h
i(1), i i(1) )
APPEND FROM datos_50s_a2_form_ord_F1.xls TYPE xl8 SHEET &h1

d1 ='c:\automatizac\datos\' +h1+'.txt'
COPY TO &d1 TYPE sdf

CREATE TABLE lisrel ( linea c(80))
APPEND FROM formMLm.LS8 TYPE sdf

GOTO 3
h2= 'datos\' +h1
instrucc = lisrel.linea
instrucc=STRTRAN(instrucc,'ç',h2)
replace lisrel.linea WITH instrucc

l1 ='c:\automatizac\ctrlisrel\' +h1+'.lpj'

COPY TO &l1 TYPE sdf

a="'C:\Archivos de programa\LISREL88\lisrel88.EXE'"

r1 ='c:\automatizac\reslisrel\' +h1+'.out'

RUN/N&a &l1 &r1

WAIT TIMEOUT 0.001
?i
ENDFOR
```

**Anexo N.** Descripción de la sintaxis Visual FoxPro, para la ejecución automatizada de análisis PLS-PM con R, modelo con variables manifiestas formativas y reflectivas, más especificado.

1.- Se borra la ventana de salida y se cierran archivos abiertos.

```
CLEAR  
CLOSE ALL
```

2.- Se desactiva la solicitud de confirmación para modificar archivos.

```
SET SAFETY OFF
```

3.- Se establece la ruta del directorio de trabajo.

```
SET PATH TO c:\automatizac
```

4.- Se inicia un ciclo “for”, determinándose la cantidad de ciclos a realizar.

```
FOR i = 1 TO 500
```

5.- Se crea la variable “h1”, la cual tomará los valores de “i” definidos en el punto 4.

```
h1 = LTRIM(STR(i))
```

6.- Se agrega la palabra “Hoja” a los valores que tomará la variable “h1”.

```
h1= 'Hoja'+h1
```

7.- Se crea una tabla Excel con 9 columnas de 5 caracteres y tantas filas como ciclos se hayan definido.

```
CREATE TABLE excel (a i(5), b i(5), c i(5), d i(5), e i(5), f i(5), g i(5),  
h i(5), i i(5) )
```

8.- Se especifica el nombre del archivo del cual se extraen los datos para ser añadidos a la tabla Excel creada en el paso anterior. Sustituyendo el nombre de cada hoja por el valor de “h1”.

```
APPEND FROM datos_500s_a0_matco.xls TYPE xl8 SHEET &h1
```

9.- Se crea la variable “d1”, que recibe el nombre de los archivos de datos con formato de texto.

```
d1 ='c:\automatizac\datos\' +h1+'.txt'
```

10.- Se copian los datos en archivos de texto creados en paso 9.

```
COPY TO &d1 TYPE sdf
```

11.- Se crea una tabla denominada “R” con 85 caracteres, que recibirá la sintaxis de R.

```
CREATE TABLE R ( linea c(85))
```

12.- Se añade la sintaxis R, contenida en el archive especificado, a la tabla Excel creada en el paso 11.

```
APPEND FROM plspm_for_w TYPE sdf
```

- 
- 13.- Se posiciona el control de ejecución en la línea nº 2 de la Tabla con la sintaxis de R (en la que se especifica la ruta para acceder a los datos brutos, el nombre de archivo es definido por “ç”).

```
GOTO 2
```

- 14.- Se crea la variable “h2”, con el contenido especificado (que es parte de la ruta para acceder a los datos).

```
h2= 'datos\\'+h1
```

- 15.- Se crea la variable “instrucc”, que contendrá una línea de comandos de R.

```
instrucc = R.linea
```

- 16.- Se cambia el nombre del archivo que contiene los datos (“ç”) por el contenido de la variable “h2”.

```
instrucc=STRTRAN(instrucc,'ç',h2)
```

- 17.- Se reemplaza el contenido de la línea nº 2 del archivo de sintaxis R por el contenido de “instrucc” (con el nombre del archivo de datos).

```
replace R.linea WITH instrucc
```

- 18.- Se posiciona el control de ejecución en la línea nº 7 de la Tabla con la sintaxis de R (en la que se especifica la ruta para guardar los resultados, el nombre de archivo es definido por “ç”).

```
GOTO 7
```

- 19.- Se crea la variable “h3”, con el contenido especificado.

```
h3= 'resR\\'+h1
```

- 20.- Se crea la variable “instrucc2”, que recibe el contenido de la línea nº 7.

```
instrucc2 = R.linea
```

- 21.- Se cambia el nombre del archivo que en el que se guardarán los resultados (“ç”) por el contenido de la variable “h3”.

```
instrucc2=STRTRAN(instrucc2,'ç',h3)
```

- 22.- Se reemplaza el contenido de la línea nº 7 del archivo de sintaxis R por el contenido de “instrucc2” (con el nombre y ruta del archivo de resultados).

```
replace R.linea WITH instrucc2
```

- 23.- Se crea la variable “l1”, que recibe la ruta con el nombre del archivo R que tiene las líneas 2 y 7 reemplazadas.

```
l1 ='c:\automatizac\ctrR\'+h1+'.r'
```

- 24.- Se graba el archivo R con el nombre archivo definido en el paso 23 en la variable “l1”.

```
COPY TO &l1 TYPE sdf
```

25.- Se crea la variable “a”, que recibe la ruta con el nombre del programa batch de R.

```
a="C:\Archivos de programa\R\R-2.14.0\bin\i386\Rscript.exe"
```

26.- Se crea la variable “r1”, que recibe la ruta con el nombre del archivo para guardar resultados.

```
r1 = 'c:\automatizac\resR\' + h1 + '.txt'
```

27.- Se ejecuta R con la sintaxis definida por “l1” y guardando los resultados en “r1”.

```
RUN/N&a &l1 &r1
```

28.- Se define el tiempo de pausa para continuar con el siguiente ciclo.

```
WAIT TIMEOUT 0.80
```

29.- Se muestra en pantalla el valor de “i”, correspondiente al ciclo “for” que se está ejecutando.

```
?i
```

30.- Se finaliza el ciclo “for”.

```
ENDFOR
```

Sintaxis Visual FoxPro:

```
CLEAR
CLOSE ALL
SET SAFETY OFF
SET PATH TO c:\automatizac

FOR i = 1 TO 500

h1 = LTRIM(STR(i))
h1= 'Hoja'+h1

CREATE TABLE excel (a i(5), b i(5), c i(5), d i(5), e i(5), f i(5), g i(5), h
i(5), i i(5) )

APPEND FROM datos_500s_a0_matco.xls TYPE XL8 SHEET &h1

d1 = 'c:\automatizac\datos\' + h1 + '.txt'

COPY TO &d1 TYPE sdf

CREATE TABLE R ( linea c(85))

APPEND FROM plspm_for_w TYPE sdf

GOTO 2
h2= 'datos\\'+h1
instrucc = R.linea
instrucc=STRTRAN(instrucc,'ç',h2)
replace R.linea WITH instrucc

GOTO 7
h3= 'resR\\'+h1
```



```
instrucc2 = R.linea
instrucc2=STRTRAN(instrucc2,'ç',h3)
replace R.linea WITH instrucc2

l1 ='c:\automatizac\ctrR\'+'h1+'.r'

COPY TO &l1 TYPE sdf

a='"C:\Archivos de programa\R\R-2.14.0\bin\i386\Rscript.exe"'

r1 ='c:\automatizac\resR\'+'h1+'.txt'

RUN/N &a &l1 &r1

WAIT TIMEOUT 0.01
?i
ENDFOR
```

**Anexo O.** Descripción de la sintaxis Visual FoxPro, para la extracción automatizada de información desde archivos de resultados de LISREL, al analizar el modelo con variables manifiestas formativas y reflectivas.

1.- Se determina el punto (".") como símbolo decimal (considerando que LISREL utiliza el punto para separar los enteros de los decimales).

```
SET POINT TO '.'
```

2.- Se define la cantidad de decimales a utilizar.

```
SET DECIMALS TO 6
```

3.- Se limpia la pantalla de salida de VFP y se cierran las tablas y bases de datos que se encuentren abiertas.

```
CLEAR  
CLOSE tables
```

4.- Se desactiva la solicitud de confirmación ante la modificación de archivos.

```
SET SAFETY OFF
```

5.- Se define la ruta del directorio de trabajo.

```
SET PATH TO c:\automatizac\reslisrel
```

6.- Se crea una tabla con los nombres de las variables que se extraerán de los archivos de resultados.

```
CREATE TABLE result1 (muestra i(4), LV1 n(8,2), LV2 f(8,2), LV3 f(8,2), ;  
LV4 f(8,2), LV5 f(8,2), LV6 f(8,2), LV7 f(8,2), ;  
LV8 f(8,2), LV9 f(8,2), BETA f(8,2), GAMMA f(8,2), ;  
RMSEA f(8,6), CFI f(8,3), SRMR f(8,3);  
)
```

7.- Se inicia un ciclo "for", determinándose la cantidad de ciclos a realizar.

```
FOR i = 1 TO 100
```

8.- Se crea la variable "h1", la cual tomará los valores de "i" definidos en el paso 7.

```
h1 = LTRIM(STR(i))
```

9.- Se agregan los nombres de los archivos de texto con los resultados LISREL a la variable "h1".

```
h1= 'Hoja'+h1+'.out'
```

10.- Se crea una tabla Excel con 7 columnas para recibir los resultados de LISREL.

```
CREATE TABLE excel (a c(14), b c(6), c c(5), d c(6), e c(27), f c(20), g  
c(7))
```

11.- Se añaden los contenidos de los archivos de resultados de LISREL a la tabla creada en el paso 10.

```
APPEND FROM &h1 TYPE SDF
```

- 12.- Se posiciona el control de ejecución a la línea 253 de la Tabla con los resultados de LISREL y se selecciona el contenido de los 14 primeros caracteres de la columna “a”.

```
GOTO 253
A = RIGHT(excel.a,14)
```

13. Se inicia un comando “if” para comprobar si existe una advertencia (asociada a PHI definida no positiva) en la línea nº 1. De existir dicha advertencia se ejecutarán los comandos que continúan al comando “then”, que seleccionan el contenido de las variables RMSEA, CFI, SRMR, LV1, LV2, LV3, LV4, LV5, LV6, LV7, LV8, LV9, BETA y GAMMA.

```
IF A = ' W_A_R_N_I_N_G' THEN
  GOTO 287
  _RMSEA = VAL( right(excel.g,5) )
  GOTO 307
  _CFI = VAL( right(excel.f,9) )
  GOTO 315
  _SRMR = VAL( right(excel.f,14) )
  GOTO 328
  _LV1 = VAL(excel.d)
  GOTO 329
  _LV2 = VAL(excel.d)
  GOTO 330
  _LV3 = VAL(excel.d)
  GOTO 331
  _LV4 = VAL(excel.e)
  GOTO 332
  _LV5 = VAL(excel.e)
  GOTO 333
  _LV6 = VAL(excel.e)
  GOTO 348
  _GAMMA = VAL(excel.b)
```

14. En el caso de no existir la advertencia en la línea nº 253 se ejecutarán los comandos ubicados luego del comando “else”, que seleccionan el contenido de las variables RMSEA, CFI, SRMR, LV1, LV2, LV3, LV4, LV5, LV6, LV7, LV8, LV9, BETA y GAMMA.

```
ELSE
  GOTO 285
  _RMSEA = VAL( right(excel.g,5) )
  GOTO 305
  _CFI = VAL( right(excel.f,9) )
  GOTO 313
  _SRMR = VAL( right(excel.f,14) )
  GOTO 326
  _LV1 = VAL(excel.d)
  GOTO 327
  _LV2 = VAL(excel.d)
  GOTO 328
  _LV3 = VAL(excel.d)
  GOTO 329
  _LV4 = VAL(excel.e)
  GOTO 330
  _LV5 = VAL(excel.e)
  GOTO 331
  _LV6 = VAL(excel.e)
  GOTO 346
  _GAMMA = VAL(excel.b)
  GOTO 347
  _BETA = VAL(excel.d)
```

```
GOTO 353
_LV7 = VAL(excel.b)
GOTO 353
_LV8 = VAL(RIGHT(excel.d,5))
GOTO 353
_LV9 = VAL(RIGHT(excel.e,5))
```

15.- Se finaliza el comando “if” iniciado en el paso 13.

```
ENDIF
```

16 - Se cierran las tablas activas de VFP.

```
CLOSE TABLES
```

17.- Se activa la Tabla que recibirá los resultados.

```
USE result1.dbf
```

18.- Se agrega un registro en blanco al final de la Tabla activa, para poder agregar la información relevante seleccionada a la Tabla activa.

```
APPEND BLANK
```

19.- Se reemplaza el contenido del campo (columna) “muestra”, por el valor de la variable “i”.

```
replace result1.muestra WITH i
```

20.- Se reemplaza el contenido de los campos correspondientes de la Tabla activa, por el valor de cada variable.

```
replace result1.lv1 WITH _LV1
replace result1.lv2 WITH _LV2
replace result1.lv3 WITH _LV3
replace result1.lv4 WITH _LV4
replace result1.lv5 WITH _LV5
replace result1.lv6 WITH _LV6
replace result1.lv7 WITH _LV7
replace result1.lv8 WITH _LV8
replace result1.lv9 WITH _LV9
replace result1.beta WITH _BETA
replace result1.gamma WITH _GAMMA
replace result1.RMSEA WITH _RMSEA
replace result1.CFI WITH _CFI
replace result1.SRMR WITH _SRMR
```

21.- Se finaliza el ciclo “for” iniciado en el paso 7.

```
ENDFOR
```

22.- Se muestra en pantalla la última Tabla activa (result1.dbf)

```
BROWSE
```

23.- Se copia el contenido de la Tabla activa a un archivo Excel.

```
COPY TO res_lisrel_50s_a0_form_ML_w.xls TYPE XL5
```

Sintaxis Visual FoxPro:

```

SET POINT TO '.'
SET DECIMALS TO 6
CLEAR
CLOSE tables
SET SAFETY OFF
SET PATH TO c:\automatizac\reslisrel

CREATE TABLE result1 (muestra i(4), LV1 n(8,2), LV2 f(8,2), LV3 f(8,2), ;
LV4 f(8,2), LV5 f(8,2), LV6 f(8,2), LV7 f(8,2), ;
LV8 f(8,2), LV9 f(8,2), BETA f(8,2), GAMMA f(8,2), ;
RMSEA f(8,3), CFI f(8,2), SRMR f(8,3);
)

FOR i = 1 TO 100
h1 = LTRIM(STR(i))
h1 = 'Hoja'+h1+'.out'

CREATE TABLE excel (a c(14), b c(6), c c(5), d c(6), e c(27), f c(20), g c(7))

APPEND FROM &h1 TYPE SDF

GOTO 253
A = RIGHT(excel.a,14)
IF A = 'W_A_R_N_I_N_G' THEN
  GOTO 287
  _RMSEA = VAL( right(excel.g,5) )
  GOTO 307
  _CFI = VAL( right(excel.f,9) )
  GOTO 315
  _SRMR = VAL( right(excel.f,14) )
  GOTO 328
  _LV1 = VAL(excel.d)
  GOTO 329
  _LV2 = VAL(excel.d)
  GOTO 330
  _LV3 = VAL(excel.d)
  GOTO 331
  _LV4 = VAL(excel.e)
  GOTO 332
  _LV5 = VAL(excel.e)
  GOTO 333
  _LV6 = VAL(excel.e)
  GOTO 348
  _GAMMA = VAL(excel.b)
  GOTO 349
  _BETA = VAL(excel.d)
  GOTO 355
  _LV7 = VAL(excel.b)
  GOTO 355
  _LV8 = VAL(RIGHT(excel.d,5))
  GOTO 355
  _LV9 = VAL(RIGHT(excel.e,5))
ELSE
  GOTO 285
  _RMSEA = VAL( right(excel.g,5) )
  GOTO 305
  _CFI = VAL( right(excel.f,9) )
  GOTO 313
  _SRMR = VAL( right(excel.f,14) )
  GOTO 326
  _LV1 = VAL(excel.d)
  GOTO 327
  _LV2 = VAL(excel.d)
  GOTO 328
  _LV3 = VAL(excel.d)
  GOTO 329
  _LV4 = VAL(excel.e)
  GOTO 330

```

```
_LV5 = VAL(excel.e)
GOTO 331
_LV6 = VAL(excel.e)
GOTO 346
_GAMMA = VAL(excel.b)
GOTO 347
_BETA = VAL(excel.d)
GOTO 353
_LV7 = VAL(excel.b)
GOTO 353
_LV8 = VAL(RIGHT(excel.d,5))
GOTO 353
_LV9 = VAL(RIGHT(excel.e,5))
ENDIF

CLOSE TABLES

USE result1.dbf
APPEND BLANK

replace result1.muestra WITH i
replace result1.lv1 WITH _LV1
replace result1.lv2 WITH _LV2
replace result1.lv3 WITH _LV3
replace result1.lv4 WITH _LV4
replace result1.lv5 WITH _LV5
replace result1.lv6 WITH _LV6
replace result1.lv7 WITH _LV7
replace result1.lv8 WITH _LV8
replace result1.lv9 WITH _LV9
replace result1.beta WITH _BETA
replace result1.gamma WITH _GAMMA
replace result1.RMSEA WITH _RMSEA
replace result1.CFI WITH _CFI
replace result1.SRMR WITH _SRMR

ENDFOR

BROWSE

COPY TO res_lisrel_50s_a0_form_ML_w.xls TYPE XL5
```

---

**Anexo P.** Descripción de la sintaxis Visual FoxPro, para la extracción automatizada de información desde archivos de resultados de R, al analizar el modelo con variables manifiestas formativas y reflectivas.

1.- Se determina el punto (“.”) como símbolo decimal (considerando que R utiliza el punto para separar los enteros de los decimales).

```
SET POINT TO '.'
```

2.- Se define la cantidad de decimales a utilizar.

```
SET DECIMALS TO 6
```

3.- Se limpia la pantalla de salida de VFP y se cierran las tablas y bases de datos que se encuentren abiertas.

```
CLEAR
CLOSE tables
```

4.- Se desactiva la solicitud de confirmación ante la modificación de archivos.

```
SET SAFETY OFF
```

5.- Se define la ruta del directorio de trabajo.

```
SET PATH TO C:\Automatizac\resR
```

6.- Se crea una tabla con los nombres de las variables que se extraerán de los archivos de resultados.

```
CREATE TABLE result1 (muestra i(4), LV1 n(8,2), LV2 f(8,2), LV3 f(8,2), ;
LV4 f(8,2), LV5 f(8,2), LV6 f(8,2), LV7 f(8,2), ;
LV8 f(8,2), LV9 f(8,2), BETA f(8,2), GAMMA f(8,2), ;
GoF f(8,3);
)
```

7.- Se inicia un ciclo “for”, determinándose la cantidad de ciclos a realizar.

```
FOR i = 1 TO 100
```

8.- Se crea la variable “h1”, la cual tomará los valores de “i” definidos en el punto 7.

```
h1 = LTRIM(STR(i))
```

9.- Se agregan los nombres de los archivos de texto con los resultados R a la variable “h1”.

```
h1 = 'Hoja'+h1+'.txt'
```

10.- Se crea una tabla Excel con 7 columnas para recibir los resultados de R.

```
CREATE TABLE excel (a c(8), b c(10), c c(10), d c(9), e c(27), f c(20), g
c(7))
```

11.- Se añaden los contenidos de los archivos de resultados de R a la tabla creada en el paso 10.

```
APPEND FROM &h1 TYPE SDF
```

- 12.- Se posiciona el control de ejecución a la línea 1 de la Tabla con los resultados de R y se borran las 35 líneas siguientes.

```
GOTO 1  
DELETE NEXT 35  
pack
```

- 13.- Se posiciona el control de ejecución en la línea nº 1 de la Tabla de resultados y se asigna el valor numérico de la columna “c” de la Tabla, a la variable “\_LV7”.

```
GOTO 1  
_LV7 = VAL(excel.c)
```

- 14.- Se asignan los valores numéricos, de las columnas correspondientes de la Tabla de resultados, a cada variable.

```
GOTO 2  
_LV8 = VAL(excel.c)  
GOTO 3  
_LV9 = VAL(excel.c)  
GOTO 5  
_LV1 = VAL(excel.c)  
GOTO 6  
_LV2 = VAL(excel.c)  
GOTO 7  
_LV3 = VAL(excel.c)  
GOTO 9  
_LV4 = VAL(excel.c)  
GOTO 10  
_LV5 = VAL(excel.c)  
GOTO 11  
_LV6 = VAL(excel.c)  
GOTO 35  
_GAMMA = VAL(RIGHT(excel.b,5))  
GOTO 41  
_BETA = VAL(RIGHT(excel.b,5))  
GOTO 61  
_GoF = VAL(RIGHT(excel.b,4))
```

- 15 - Se cierran las tablas activas de VFP.

```
CLOSE TABLES
```

- 16.- Se activa la Tabla que recibirá los resultados.

```
USE result1.dbf
```

- 17.- Se agrega un registro en blanco al final de la Tabla activa, para poder agregar la información relevante seleccionada a la Tabla activa.

```
APPEND BLANK
```

- 18.- Se reemplaza el contenido del campo (columna) “muestra”, por el valor de la variable “i”.

```
replace result1.muestra WITH i
```



19.- Se reemplaza el contenido de los campos correspondientes de la Tabla activa, por el valor de cada variable.

```
replace result1.lv1 WITH _LV1
replace result1.lv2 WITH _LV2
replace result1.lv3 WITH _LV3
replace result1.lv4 WITH _LV4
replace result1.lv5 WITH _LV5
replace result1.lv6 WITH _LV6
replace result1.lv7 WITH _LV7
replace result1.lv8 WITH _LV8
replace result1.lv9 WITH _LV9
replace result1.beta WITH _BETA
replace result1.gamma WITH _GAMMA
replace result1.gof WITH _GoF
```

20.- Se finaliza el ciclo "for".

```
ENDFOR
```

21.- Se muestra en pantalla la última Tabla activa (result1.dbf)

```
BROWSE
```

22.- Se copia el contenido de la Tabla activa a un archivo Excel.

```
COPY TO res_R_50s_a0_form_plspm_w.xls TYPE XL5
```

Sintaxis Visual FoxPro:

```
SET POINT TO '.'
SET DECIMALS TO 6
CLEAR
CLOSE tables
SET SAFETY OFF
SET PATH TO C:\Automatizac\resR

CREATE TABLE result1 (muestra i(4), LV1 n(8,2), LV2 f(8,2), LV3 f(8,2), ;
LV4 f(8,2), LV5 f(8,2), LV6 f(8,2), LV7 f(8,2), ;
LV8 f(8,2), LV9 f(8,2), BETA f(8,2), GAMMA f(8,2), ;
GoF f(8,2);
)

FOR i = 1 TO 100

h1 = LTRIM(STR(i))
h1 = 'Hoja'+h1+'.txt'

CREATE TABLE excel (a c(8), b c(10), c c(10), d c(9), e c(27), f c(20), g
c(7))

APPEND FROM &h1 TYPE SDF

GOTO 1
DELETE NEXT 35
pack

GOTO 1
_LV7 = VAL(excel.c)
GOTO 2
_LV8 = VAL(excel.c)
GOTO 3
_LV9 = VAL(excel.c)
```

```
GOTO 5
_LV1 = VAL(excel.c)
GOTO 6
_LV2 = VAL(excel.c)
GOTO 7
_LV3 = VAL(excel.c)
GOTO 9
_LV4 = VAL(excel.c)
GOTO 10
_LV5 = VAL(excel.c)
GOTO 11
_LV6 = VAL(excel.c)
GOTO 35
_GAMMA = VAL(RIGHT(excel.b,5))
GOTO 41
_BETA = VAL(RIGHT(excel.b,5))
GOTO 61
_GoF = VAL(RIGHT(excel.b,4))

CLOSE TABLES

USE result1.dbf
APPEND BLANK

replace result1.muestra WITH i
replace result1.lv1 WITH _LV1
replace result1.lv2 WITH _LV2
replace result1.lv3 WITH _LV3
replace result1.lv4 WITH _LV4
replace result1.lv5 WITH _LV5
replace result1.lv6 WITH _LV6
replace result1.lv7 WITH _LV7
replace result1.lv8 WITH _LV8
replace result1.lv9 WITH _LV9
replace result1.beta WITH _BETA
replace result1.gamma WITH _GAMMA
replace result1.gof WITH _GoF

ENDFOR

BROWSE

COPY TO res_R_50s_a0_form_plspm_w.xls TYPE XL5
```

---

**Anexo Q.** Descripción de la sintaxis Visual FoxPro, para la extracción automatizada de información desde archivos de resultados de GeSCA, al analizar el modelo con variables manifiestas formativas y reflectivas.

1.- Se determina el punto (".") como símbolo decimal.

```
SET POINT TO '.'
```

2.- Se define la cantidad de decimales a utilizar.

```
SET DECIMALS TO 6
```

3.- Se limpia la pantalla de salida de VFP y se cierran las tablas y bases de datos que se encuentren abiertas.

```
CLEAR  
CLOSE tables
```

4.- Se desactiva la solicitud de confirmación ante la modificación de archivos.

```
SET SAFETY OFF
```

5.- Se define la ruta del directorio de trabajo.

```
SET PATH TO C:\Automatizac\resGeSCA
```

6.- Se crea una tabla con los nombres de las variables que se extraerán de los archivos de resultados.

```
CREATE TABLE result1 (muestra i(4), LV1 n(8,2), LV2 f(8,2), LV3 f(8,2), ;  
LV4 f(8,2), LV5 f(8,2), LV6 f(8,2), LV7 f(8,2), ;  
LV8 f(8,2), LV9 f(8,2), BETA f(8,2), GAMMA f(8,2);  
)
```

7.- Se inicia un ciclo "for", determinándose la cantidad de ciclos a realizar.

```
FOR i = 1 TO 100
```

8.- Se crea la variable "h1", la cual tomará los valores de "i" definidos en el punto 7.

```
h1 = LTRIM(STR(i))
```

9.- Se agregan los nombres de los archivos de texto con los resultados GeSCA a la variable "h1".

```
h1 = 'Hoja'+h1+'.txt'
```

10.- Se crea una tabla Excel con 7 columnas para recibir los resultados de GeSCA.

```
CREATE TABLE excel (a c(8), b c(10), c c(10), d c(9), e c(27), f c(20), g  
c(7))
```

11.- Se añaden los contenidos de los archivos de resultados de R a la tabla creada en el paso 10.

```
APPEND FROM &h1 TYPE SDF
```

- 12.- Se posiciona el control de ejecución en la línea nº29 de la Tabla de resultados y se asigna el valor numérico de la columna “a” de la Tabla, a la variable “\_LV1.

```
GOTO 29
_LV1 = VAL(excel.a)
```

- 13.- Se asignan los valores numéricos, de las columnas correspondientes de la Tabla de resultados, a cada variable.

```
GOTO 39
_LV2 = VAL(excel.a)
GOTO 49
_LV3 = VAL(excel.a)
GOTO 62
_LV4 = VAL(excel.a)
GOTO 72
_LV5 = VAL(excel.a)
GOTO 82
_LV6 = VAL(excel.a)
GOTO 98
_LV7 = VAL(excel.a)
GOTO 108
_LV8 = VAL(excel.a)
GOTO 118
_LV9 = VAL(excel.a)
GOTO 131
_BETA = VAL(excel.a)
GOTO 135
_GAMMA = VAL(excel.a)
```

- 14 - Se cierran las tablas activas de VFP.

```
CLOSE TABLES
```

- 15.- Se activa la Tabla que recibirá los resultados.

```
USE result1.dbf
```

- 16.- Se agrega un registro en blanco al final de la Tabla activa, para poder agregar la información relevante seleccionada a la Tabla activa.

```
APPEND BLANK
```

- 17.- Se reemplaza el contenido del campo (columna) “muestra”, por el valor de la variable “i”.

```
replace result1.muestra WITH i
```

- 18.- Se reemplaza el contenido de los campos correspondientes de la Tabla activa, por el valor de cada variable.

```
replace result1.lv1 WITH _LV1
replace result1.lv2 WITH _LV2
replace result1.lv3 WITH _LV3
replace result1.lv4 WITH _LV4
replace result1.lv5 WITH _LV5
replace result1.lv6 WITH _LV6
replace result1.lv7 WITH _LV7
replace result1.lv8 WITH _LV8
replace result1.lv9 WITH _LV9
replace result1.beta WITH _BETA
replace result1.gamma WITH _GAMMA
```

19.- Se finaliza el ciclo "for".

```
ENDFOR
```

20.- Se muestra en pantalla la última Tabla activa (result1.dbf)

```
BROWSE
```

21.- Se copia el contenido de la Tabla activa a un archivo Excel.

```
COPY TO res_gesca_50s_a0_form_w.xls TYPE XL5
```

Sintaxis Visual FoxPro:

```
SET POINT TO '.'
SET DECIMALS TO 6
CLEAR
CLOSE tables
SET SAFETY OFF
SET PATH TO c:\automatizac\resGeSCA

CREATE TABLE result1 (muestra i(4), LV1 n(8,3), LV2 f(8,3), LV3 f(8,3), ;
LV4 f(8,3), LV5 f(8,3), LV6 f(8,3), LV7 f(8,3), ;
LV8 f(8,3), LV9 f(8,3), BETA f(8,3), GAMMA f(8,3);
)

FOR i = 1 TO 100

h1 = LTRIM(STR(i))
h1 = 'Hoja'+h1+'.txt'
CREATE TABLE excel (a c(14), b c(6), c c(5), d c(6), e c(27), f c(20), g c(7))

APPEND FROM &h1 TYPE SDF

GOTO 29
  _LV1 = VAL(excel.a)
GOTO 39
  _LV2 = VAL(excel.a)
GOTO 49
  _LV3 = VAL(excel.a)
GOTO 62
  _LV4 = VAL(excel.a)
GOTO 72
  _LV5 = VAL(excel.a)
GOTO 82
  _LV6 = VAL(excel.a)
GOTO 98
  _LV7 = VAL(excel.a)
GOTO 108
  _LV8 = VAL(excel.a)
GOTO 118
  _LV9 = VAL(excel.a)
GOTO 131
  _BETA = VAL(excel.a)
GOTO 135
  _GAMMA = VAL(excel.a)

CLOSE TABLES

USE result1.dbf
APPEND BLANK

replace result1.muestra WITH i
replace result1.lv1 WITH _LV1
```

## Anexos

---

```
replace result1.lv2 WITH _LV2
replace result1.lv3 WITH _LV3
replace result1.lv4 WITH _LV4
replace result1.lv5 WITH _LV5
replace result1.lv6 WITH _LV6
replace result1.lv7 WITH _LV7
replace result1.lv8 WITH _LV8
replace result1.lv9 WITH _LV9
replace result1.beta WITH _BETA
replace result1.gamma WITH _GAMMA

ENDFOR
BROWSE
COPY TO res_gesca_50s_a0_form_w.xls TYPE XL5
```