

4

VALORACIÓN DE LOS RESULTADOS EN PLS-SEM 1ª PARTE

Evaluación del modelo de medida reflectivo

EJEMPLO ILUSTRATIVO: MODELOS DE MEDIDA REFLECTIVOS

Ejecutar el algoritmo PLS-SEM

Continuamos trabajando con nuestro ejemplo del modelo PLS-SEM sobre reputación corporativa. En el Capítulo 3, se explicaba cómo estimar el modelo PLS y cómo obtener los resultados al abrir el informe predeterminado en el software SmartPLS 4 (Ringle, Wende y Becker, 2022). Insistimos que para hacer esto, se debe en primer lugar cargar el modelo simple de reputación corporativa y tras ello correr el modelo pinchando en el icono de la rueda dentada denominado *Calcular* en la barra de herramientas o utilizando el menú desplegable yendo a *Calcular* → *Algoritmo PLS-SEM*. Después de ejecutar el *Algoritmo PLS-SEM*, si previamente ha marcado *Abrir informe* en el cuadro de diálogo, se abrirá automáticamente el informe de resultados de SmartPLS; si no es así, haga clic en el botón *Abrir informe* en la ventana de *Modelización*. Para guardar los resultados de este informe pinche en el botón *Guardar* de la barra de herramientas.

Antes de analizar los resultados, se necesita comprobar de forma rápida si el algoritmo convergió (*i.e.*, si se alcanzó el criterio de parada del algoritmo y no se llegó al número máximo de iteraciones). Para hacer esto, vaya a *Algoritmo* → *Cambios en el criterio de parada* en el informe de resultados. Entonces aparece la tabla que se muestra en la Figura A4.1, que indica el número de iteraciones del algoritmo PLS-SEM. Este número debería ser siempre menor que el número máximo de iteraciones (p. ej., 3.000). En la parte inferior izquierda de la tabla, se puede observar que el algoritmo alcanzó la convergencia tras la 5ª iteración.

Si el algoritmo PLS-SEM no converge en menos de 3.000 iteraciones (el ajuste por defecto incluido en el software), quiere decir que el algoritmo no pudo encontrar una solución estable. Esta situación rara vez ocurre. Pero si sucediera, existen problemas con los datos y éstos necesitan de un análisis minucioso. Por ejemplo, pueden surgir problemas con los datos cuando la muestra es demasiado pequeña o uno de los indicadores tiene muchos valores idénticos (*i.e.*, las mismas puntuaciones en un mismo indicador, que impiden que haya una variabilidad suficiente).

Figura A.4.1 ■ Tabla del criterio de parada en SmartPLS

	comp_1	comp_2	comp_3	cusa	cusl_1	cusl_2	cusl_3	like_1	like_2	like_3
Iteración 0	0,401	0,401	0,401	1,000	0,386	0,386	0,386	0,386	0,386	0,386
Iteración 1	0,536	0,341	0,328	1,000	0,368	0,421	0,365	0,419	0,378	0,359
Iteración 2	0,536	0,340	0,328	1,000	0,369	0,420	0,365	0,418	0,378	0,360
Iteración 3	0,536	0,340	0,328	1,000	0,369	0,420	0,365	0,418	0,378	0,360
Iteración 4	0,536	0,340	0,328	1,000	0,369	0,420	0,365	0,418	0,378	0,360
Iteración 5	0,536	0,340	0,328	1,000	0,369	0,420	0,365	0,418	0,378	0,360

Cuando la estimación del modelo PLS converge, algo que prácticamente ocurre siempre, es necesario examinar las siguientes tablas de resultados de los cálculos de PLS-SEM en el informe de resultados relativo a la evaluación del modelo de medida reflectivo: *Cargas externas*, *Fiabilidad compuesta (rho_a)*, *Fiabilidad compuesta (rho_c)*, *Alfa de Cronbach*, *Varianza extraída media (AVE)* y *Validez discriminante*. Examinaremos en los Capítulos 5 y 6 otros elementos que aparecen en el informe de resultados, cuando se amplía el actual modelo simple, incluyendo medidas formativas, y se valoran los resultados del modelo estructural.

Evaluación del modelo de medida reflectivo

El modelo simple de reputación corporativa consta de tres variables latentes con modelos de medida reflectivos (*i.e.*, *COMP*, *CUSL* y *LIKE*) junto a un constructo con un único indicador (*CUSA*). Para los modelos de medida reflectivos, se necesitan las estimaciones de las relaciones entre los constructos reflectivos y sus indicadores (*i.e.*, cargas externas). La Figura A4.2 muestra la tabla de resultados con las cargas externas, las cuales se puede encontrar en *Resultados finales* → *Cargas*. Por defecto, las cargas externas también se muestran en la *Salida gráfica* del informe de resultados de PLS tras correr el algoritmo PLS-SEM. Todas las cargas externas de las constructos reflectivos *COMP*, *CUSL* y *LIKE* están muy por encima del valor mínimo de 0,708, lo que sugiere niveles suficientes de fiabilidad del indicador. El indicador *comp_2* (carga externa: 0,798) tiene la menor fiabilidad del ítem con un valor de 0,637 (0,798²), mientras que el indicador *cusl_2* (carga externa: 0,917) tiene la mayor fiabilidad del ítem, con un valor de 0,841 (0,917²).

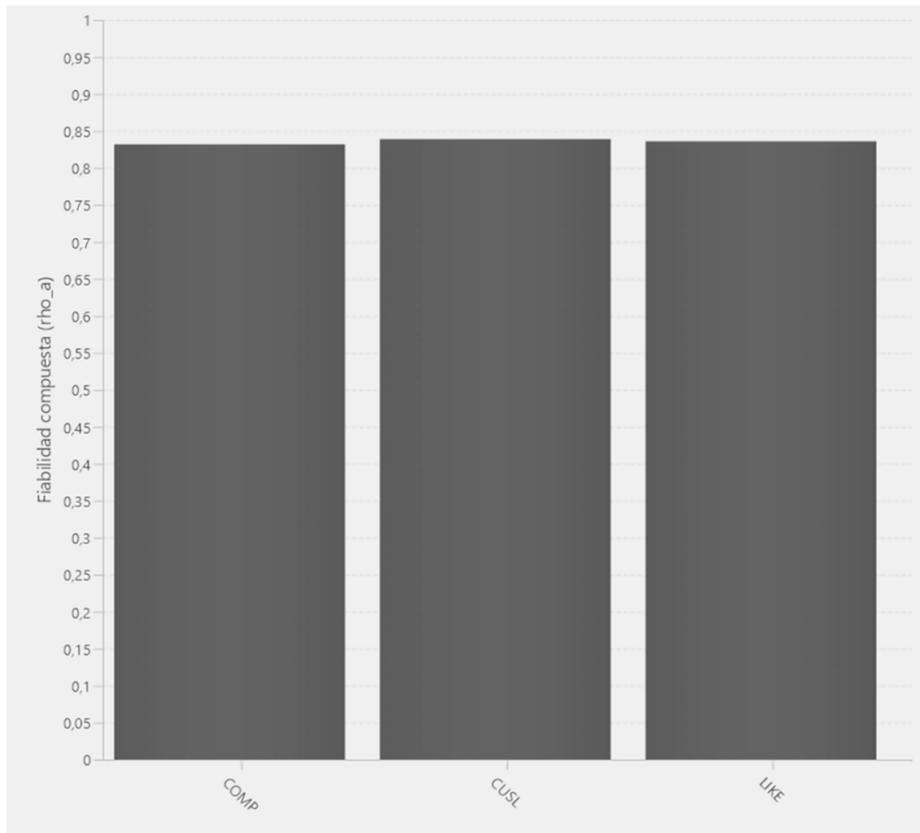
Figura A.4.2 ■ Cargas externas

	COMP	CUSA	CUSL	LIKE
comp_1	0,858			
comp_2	0,798			
comp_3	0,818			
cusa		1,000		
cusl_1			0,833	
cusl_2			0,917	
cusl_3			0,843	
like_1				0,879
like_2				0,870
like_3				0,843

A continuación, vaya a *Criterios de calidad* → *Fiabilidad y validez* de constructo para evaluar la fiabilidad compuesta del constructo y la validez convergente. Aquí, tiene la opción de mostrar los valores de fiabilidad compuesta usando bien un gráfico de barras, bien usando una matriz. La Figura A4.3 muestra los valores de las medidas de fiabilidad de la consistencia interna en formato de matriz. Con valores ρ_A de 0,832 (*COMP*), 0,839 (*CUSL*) y 0,836 (*LIKE*), cada uno de los tres constructos presenta altos niveles de fiabilidad de consistencia interna. Pinchando en *Gráfico de barras* → *Fiabilidad compuesta (rho_a)* se muestra el gráfico de barras de los valores de fiabilidad de los constructos (Figura A4.4). Si un valor ρ_A está por encima del valor mínimo aceptable, la barra correspondiente aparece en verde. Si el valor ρ_A es inferior a 0,70, la barra se muestra en rojo. Como se ha indicado anteriormente, todos los valores ρ_A superan el umbral. Tenga en cuenta que el valor ρ_A de la variable con un solo indicador *CUSA* es 1,00, pero esto no puede interpretarse como evidencia de que la medición del constructo es perfectamente fiable y no debe ser reportada junto a otras medidas de fiabilidad.

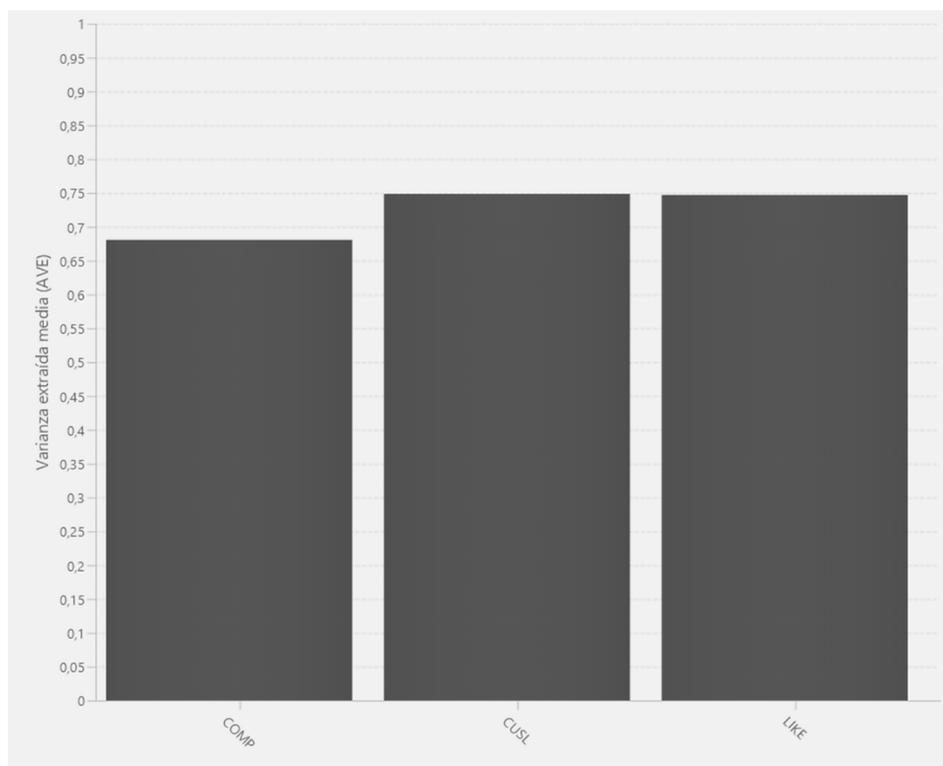
Figura A.4.3 ■ Fiabilidad y validez del constructo

	Alfa de Cronbach	Fiabilidad compuesta (rho_a)	Fiabilidad compuesta (rho_c)	Varianza extraída media (AVE)
COMP	0,776	0,832	0,865	0,681
CUSL	0,831	0,839	0,899	0,748
LIKE	0,831	0,836	0,899	0,747

Figura A.4.4 ■ Fiabilidad ρ_A 

De forma análoga a ρ_A , usted puede examinar los gráficos de barras del alfa de Cronbach y la fiabilidad compuesta ρ_C . Todas las barras del gráfico están en verde, indicando que todas las medidas para todos los constructos están por encima del mínimo aceptable de 0,70. Para ver los valores del alfa de Cronbach específicos (0,776 para *COMP*, 0,831 para *CUSL*, y 0,831 para *LIKE*) y la fiabilidad compuesta ρ_C (0,865 para *COMP*, 0,899 para *CUSL* y 0,899 para *LIKE*) vaya a *Criterios de calidad* → *Fiabilidad y validez del constructo* → *Resumen*. De nuevo, como *CUSA* se mide con un solo indicador, la interpretación de los valores del alfa de Cronbach y de la fiabilidad compuesta de este constructo no es relevante.

La evaluación de la validez convergente se basa en los valores de AVE, a los que se puede acceder navegando a *Criterios de calidad* → *Fiabilidad y validez del constructo* en el informe de resultados. Como en el caso de las medidas de fiabilidad y consistencia interna, SmartPLS ofrece la opción de mostrar los resultados mediante un gráfico de barras (Figura A4.5) o en formato de matriz. En este ejemplo, los valores AVE de *COMP* (0,681), *CUSL* (0,748) y *LIKE* (0,747) están muy por encima del nivel mínimo requerido de 0,50. Por tanto, las medidas de los tres constructos reflectivos presentan niveles altos en lo que a validez convergente se refiere.

Figura A.4.5 ■ Varianza extraída media (AVE)

Finalmente, en *Criterios de calidad* → *Validez discriminante*, SmartPLS ofrece varios enfoques para evaluar si las medidas del constructo muestran validez discriminante a nivel empírico. El criterio más fiable para la evaluación de la validez discriminante es la ratio *Heterotrait-monotrait (HTMT)*, al que se puede acceder a través de la sección de *Criterios de calidad* → *Validez discriminante* → *Matriz de ratios Heterotrait-monotrait (HTM)* del informe de resultados. La figura A4.6 muestra los valores HTMT para todos los pares de constructos en formato matriz. Cabe señalar que SmartPLS calcula los valores HTMT en base a correlaciones absolutas para evitar los resultados adversos debidos a una combinación de correlaciones de indicadores positivos y negativos (Ringle, Sarstedt, Sinkovics y Sinkovics, 2023). El informe de resultados también muestra estos valores HTMT en formato de lista y de gráfico de barras, tomando 0,85 como nivel máximo relevante. Como se puede ver, todos los valores de HTMT están claramente por debajo del nivel máximo más conservador de 0,85, incluso para *COMP* y *LIKE*, así como *CUSA* y *CUSL*, que, desde un punto de vista conceptual, son bastante similares. Recuerde que el valor máximo del HTMT para constructos conceptualmente similares es 0,90.

Según el criterio de Fornell-Larcker, la raíz cuadrada del AVE de cada constructo debería ser mayor que la correlación más alta que tiene ese constructo con cualquiera de los otros constructos del modelo (esta idea es idéntica a comparar el AVE con las correlaciones al cuadrado de los distintos constructos). La figura A4.7 muestra los resultados de la evaluación del criterio Fornell-Larcker con la raíz cuadrada del AVE de los constructos reflectivos en la diagonal y las correlaciones entre los diferentes constructos fuera de la diagonal. Así, por ejemplo, el constructo reflectivo *COMP* tiene un valor de 0,825 para la raíz cuadrada de su AVE, que ha de ser comparado con todas las correlaciones en la columna de *COMP*. Tenga presente que para *CUSL* se ha de considerar todas las correlaciones tanto por filas como por columnas y, para *LIKE*, sólo las de la fila. De forma general, las raíces cuadradas de los valores AVE para los constructos reflectivos *COMP* (0,825), *CUSL* (0,865) y *LIKE* (0,864) son todos superiores a las correlaciones de estos constructos con cualquiera de las otras variables latentes presentes en el modelo, lo que indica que todos los constructos son medidas válidas de conceptos únicos.

Tenga presente que, aunque fue usado con frecuencia en investigación aplicada, el criterio Fornell-Larcker no detecta de forma fiable problemas de validez discriminante. Por lo tanto, el criterio no es adecuado para revelar los problemas de validez discriminante existentes. No obstante, cualquier incumplimiento del criterio de Fornell-Larcker debe tomarse como evidencia sólida de un problema de validez discriminante grave.

Figura A.4.6 ■ HTMT

	COMP	CUSA	CUSL	LIKE
COMP				
CUSA	0,465			
CUSL	0,532	0,755		
LIKE	0,780	0,577	0,737	

Figura A.4.7 ■ Criterio de Fornell - Larcker

	COMP	CUSA	CUSL	LIKE
COMP	0,825			
CUSA	0,436	1,000		
CUSL	0,450	0,689	0,865	
LIKE	0,645	0,528	0,615	0,864

Además de valorar las ratios HTMT, se debería comprobar si los valores HTMT son significativamente diferentes del valor máximo. Específicamente, asumimos un valor máximo de 0,85 para todos los pares de constructos excepto *COMP* y *LIKE*, así como *CUSA* y *CUSL*, para los cuales asumimos un límite máximo más alto (0,90) debido a su similitud conceptual. Para ello se necesita calcular los intervalos de confianza *bootstrap* a través de la ejecución de la opción *bootstrapping*. Para ejecutar el procedimiento *bootstrapping*, se ha de volver a la ventana de modelización haciendo clic en el símbolo de flecha con la etiqueta *Edición*. A continuación, haga clic con el botón izquierdo en *Calcular* → *Bootstrapping* en el menú desplegable. En el cuadro de diálogo que se abre, se ha de elegir la opción de *bootstrapping*, tal como se muestra en la Figura A4.8 (el Capítulo 5 contiene una introducción más detallada al procedimiento *bootstrapping* y la configuración de parámetros). Hemos de asegurarnos de seleccionar 10.000 *Submuestras* y la opción de *Completo (más lento)*, que incluye los resultados para HTMT (a diferencia de la opción *Con lo más importante (más rápido)*). Seleccione *Percentil bootstrap* como *Método del intervalo de confianza*, *Una cola* como *Tipo de test* y un *nivel de significación* de 0,05 (tenga presente que puede seleccionar *Dos colas* como *Tipo de test* y un nivel de significación de 0,10). Se debe seleccionar *Semilla fija* en el *Generador de números aleatorios* para facilitar la reproducibilidad de los resultados. Por último, asegúrese de marcar la casilla situada al lado de *Abrir informe* y haga clic en *Iniciar cálculos*.

Figura A.4.8 ■ Opciones de bootstrapping en SmartPLS



Tras ejecutar *bootstrapping*, se abre automáticamente el informe de resultados. Vaya a *Criterios de calidad* → *Ratio Heterotrait-monotrait (HTMT)* y navegue a *Intervalos de confianza*. La tabla que se abre (Figura A4.9) muestra los valores de HTMT originales (columna de *Muestra original (O)*) para cada par combinación de constructos en el modelo, junto con los valores de HTMT medios calculados a partir de las 10.000 submuestras de *bootstrap*, tal como se muestra en la columna *Muestra original (O)*. Tenga presente que debería obtener los mismos resultados porque seleccionó la opción *Semilla fija* en los ajustes del *bootstrapping*. No obstante, si selecciona *Semilla aleatoria*, los resultados probablemente difieran de sus propios resultados y volverán a cambiar ligeramente si vuelve a ejecutar el procedimiento *bootstrapping*. El motivo es que el *bootstrapping* se basa en muestras de *bootstrap* extraídas de forma aleatoria, que serán diferentes cada vez que se ejecuta el procedimiento. Sin embargo, las diferencias en los resultados *bootstrapping* suelen ser marginales, siempre que se hayan extraído un número suficiente de submuestras *bootstrap* (p. ej., 10.000). Al seleccionar la opción *Semilla fija*, el generador de números aleatorios del algoritmo *bootstrapping* utiliza una configuración fija, lo que facilita la reproducción de los resultados en las estimaciones posteriores del modelo.

Figura A.4.9 ■ Intervalos de confianza (con sesgo corregido) para HTMT

	Muestra original (O)	Media de la muestra (M)	Sesgo	5.0%	95.0%
CUSA ↔ COMP	0,465	0,464	-0,001	0,369	0,556
CUSL ↔ COMP	0,532	0,532	-0,000	0,429	0,625
CUSL ↔ CUSA	0,755	0,754	-0,000	0,695	0,805
LIKE ↔ COMP	0,780	0,780	0,000	0,705	0,843
LIKE ↔ CUSA	0,577	0,577	-0,000	0,503	0,644
LIKE ↔ CUSL	0,737	0,737	-0,000	0,664	0,800

Las columnas etiquetadas con 5% y 95% muestran los límites inferior y superior para un intervalo de confianza *bootstrap* de una cola del 95% (o el intervalo de confianza de bootstrap de dos colas del 90%, respectivamente). La prueba estadística se centra en la cola derecha de la distribución *bootstrap* para mostrar que el valor HTMT es significativamente menor que los límites máximos correspondientes (0,85 y 0,90) con una probabilidad de error del 5%. Este es el caso, si el resultado del límite superior mostrado en la columna 95% es inferior a 0,85 y en el caso de *COMP* y *LIKE*, así como *CUSA* y *CUSL*, inferior a 0,90. Más importante aún, incluso suponiendo un límite máximo más conservador de 0,85 para todas las combinaciones de constructos (*i.e.*, incluyendo *CUSA* y *CUSL*, así como *COMP* y *LIKE*), encontramos que todos los valores HTMT son significativamente más bajos que este valor (*i.e.*, el límite superior de los intervalos de confianza es menor que 0,85). Por ejemplo, los límites inferior y superior del intervalo de confianza del 95% del HTMT para *CUSA* y *COMP* son

Figura A.4.10 ■ Resumen de los resultados para los modelos de medida reflectivos

Variable latente	Indicadores	Validez convergente			Fiabilidad de consistencia interna			Validez discriminante
		Cargas	Fiabilidad del indicador	AVE	Alfa de Cronbach	Fiabilidad ρ_A	Fiabilidad del compuesto ρ_C	
		>0,70	>0,50	>0,50	0,60-0,90	0,60-0,90	0,60-0,90	¿Significativamente inferior a 0,85 (0,90)?
COMP	comp_1	0,858	0,736	0,681	0,776	0,832	0,865	Sí
	comp_2	0,798	0,637					
	comp_3	0,818	0,669					
CUSL	cusl_1	0,833	0,694	0,748	0,831	0,839	0,899	Sí
	cusl_2	0,917	0,841					
	cusl_3	0,843	0,711					
LIKE	like_1	0,879	0,773	0,747	0,831	0,836	0,899	Sí
	like_2	0,870	0,757					
	like_3	0,843	0,711					

0,369 y 0,556, respectivamente. Dado que el límite superior de 0,556 es inferior a 0,85, el valor HTMT de 0,465 para *CUSA* y *COMP* es significativamente menor que el límite máximo más conservador de 0,85. En resumen, los resultados del intervalo de confianza calculados con bootstrap para el criterio HTMT también demuestran claramente la validez discriminante de los constructos.

La figura A4.10 resume los resultados de la evaluación del modelo de medida reflectivo. Como se puede ver, se cumplen todos los criterios de evaluación del modelo, proporcionando evidencias de la fiabilidad y validez de las medidas.