



Binh PLS-SEM Smartpls 4

quản trị kinh doanh (Trường Đại học Kinh tế Thành phố Hồ Chí Minh)



Scan to open on Studeersnel

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/378342710>

Ước lượng, đánh giá và diễn giải mô hình PLS-SEM sử dụng SmartPLS 4

Method · February 2024

CITATIONS

0

READS

1,857

1 author:



Phung Thanh Binh
UEH University
6 PUBLICATIONS 129 CITATIONS

[SEE PROFILE](#)

ĐẠI HỌC KINH TẾ TP.HCM
TRƯỜNG KINH TẾ, LUẬT VÀ QUẢN LÝ NHÀ NƯỚC
KHOA KINH TẾ

**ƯỚC LUỢNG, KIỂM ĐỊNH VÀ DIỄN GIẢI CÁC
MÔ HÌNH PLS-SEM VỚI SMARTPLS 4**

Phùng Thanh Bình - 2024
(ptbinh@ueh.edu.vn)

Từ năm 2014, mô hình **PLS-SEM** (mô hình phương trình cấu trúc bình phương tối thiểu riêng phần) được sử dụng ngày càng nhiều trong các bài báo khoa học uy tín trên thế giới, nhất là trong lĩnh vực nghiên cứu thị trường, marketing, quản trị, và tâm lý. Với ngày càng nhiều sinh viên bậc đại học chọn học các chương trình đào tạo chất lượng cao, và hiện tại là các chương trình đào tạo tiếng Anh bán phần, tiếng Anh toàn phần tại **Đại học Kinh tế TP.HCM (UEH)** thì nhu cầu nghiên cứu khoa học của sinh viên ngày càng gia tăng, trong đó mô hình PLS-SEM đã và đang là một lựa chọn khá phổ biến ở các ngành thuộc khối kinh doanh. Vài năm gần đây, tôi tham gia bình duyệt nhiều bài nghiên cứu khoa học của sinh viên **UEH** và nhận thấy rằng khá nhiều sinh viên chưa hiểu đúng mô hình PLS-SEM, cách ước lượng và nhất là đánh giá, giải thích các kết quả. Rất có thể một lượng không ít các học viên cao học cũng gặp các vấn đề tương tự. Chính vì thế, tôi biên soạn tài liệu này theo hướng thực hành trên phần mềm **SmartPLS 4** để sinh viên mạnh dạn sử dụng và áp dụng một cách đúng đắn các mô hình PLS-SEM cho các nghiên cứu của mình. Hy vọng rằng tài liệu này sẽ đóng góp một phần nào đó vào sứ mệnh nghiên cứu khoa học của **UEH** nói chung, cũng như các sinh viên **UEH** có niềm đam mê nghiên cứu khoa học hiện tại và tương lai.

1. CÀI SMARTPLS 4

Truy cập website: <https://www.smartpls.com/downloads/> (hoặc tìm kiếm SMARTPLS trên Google), chọn phiên bản phù hợp với máy tính của bạn (ví dụ: *SmartPLS 4 for Microsoft Windows*), nhấp chuột vào biểu tượng “**Download Windows Installer**”, tải về máy tính, sau đó nhấp đúp vào tệp *smartpls4_windows-x64.exe* vừa tải về và tiến hành cài đặt. Khi xuất hiện hộp thoại “**Licensing options**”, hãy chọn “**Free 30 days trial license**¹”, nhập Tên + Họ và E-mail của bạn rồi bấm **Submit**, ngay tức thì bạn sẽ nhận mã xác nhận trong E-mail và nhập vào ô “**Confirmation token**”.

¹ Nếu chọn “**Student license**” thì bạn sẽ được sử dụng miễn phí, nhưng bị hạn chế dữ liệu tối đa 100 quan sát. Điều này sẽ bất tiện với tệp dữ liệu được sử dụng trong hướng dẫn này (hơn 300 quan sát).

2. DỮ LIỆU VÀ TÀI LIỆU HƯỚNG DẪN

Truy cập: <https://www.pls-sem.net/downloads/>, chọn *3rd Edition: A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*.

Book files

- Advanced Issues in Partial Least Squares Structural Equation Modeling (PLS-SEM)
- Partial Least Squares Structural Equation Modeling (PLS-SEM) Using R
- • 3rd Edition: A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)
- 2nd Edition: A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)
- 1st Edition: A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)

Tải các tệp dữ liệu sau đây:

- **SmartPLS 4 Case Studies:** Là các tinh huống cùng một tệp dữ liệu **Corporate reputation data**, từ Chapter 2 đến Chapter 7 trong giáo trình của Hair và công sự. (2022). Sau đây gọi là “Case Study, Chapter x”. Thao tác trong các tinh huống này sẽ được trình bày lại một cách dễ hiểu trong tài liệu này.
- **Full data set (n=344):** Tải tệp dữ liệu “**Corporate reputation data.txt**” để sử dụng cho phiên bản **Professional** dùng thử 30 ngày.
- **Reduced data set (n=99):** Tải tệp dữ liệu “**Corporate reputation data.txt**” để sử dụng cho phiên bản **Student** với số quan sát tối đa là 100.

Các biến quan sát (**observed variables**)/các biến chỉ báo (**indicators**) của các biến tiềm ẩn (**latent variables**)/các cấu trúc (**constructs**) được đo lường từ nhiều biến quan sát khác nhau (multiple items) của mô hình đơn giản [**Simple model** = chỉ bao gồm các cấu trúc của mô hình đo lường kết quả = **Reflective Measurement Model** (bao gồm: **COMP** = Company's competence, **LIKE** = Company's likeability, và **CUSL** = Customer loyalty) được trình bày trong **Exhibit A2.3** (Case Study, Chapter 2: trang 4) hoặc **Exhibit 2.16** (Hair và công sự., 2022: trang 70). Các biến quan sát/chỉ báo này được đo bằng thang đo Likert 7 điểm với các tên biến là: **comp_1, comp_2, comp_3; like_1, like_2, like_3; cusl_1, cusl_2, cusl_3**. Riêng biến tiềm ẩn/cấu trúc về sự hài lòng của khách hàng (**CUSA** = Customer satisfaction) được cấu thành bởi một biến quan sát duy nhất (single item), thang đo Likert 7 điểm. Biến quan sát này được đặt tên là **cusa** trong tệp dữ liệu.

Các biến quan sát/các chỉ báo của các biến tiềm ẩn/các cấu trúc được đo lường từ nhiều biến quan sát khác nhau của mô hình mở rộng [**Extended model** = bao gồm cả các cấu trúc của mô hình đo lường kết quả như ở trên và các cấu trúc của mô hình đo lường nguyên nhân = **Formative Measurement Model** (bao gồm: **QUAL** = Quality of a company's products and services, **PERF** = Company's economic and managerial performance, **CSOR** = Company's corporate social responsibility, và **ATTR** = Company's attractiveness) được trình bày trong **Exhibit A5.2** (Case Study, Chapter 5: trang 3) hoặc **Exhibit 5.13** (Hair và công sự., 2022: trang 160-161). Các biến quan sát này được đo bằng thang đo Likert 7 điểm với tên các biến là: **qual_1** đến **qual_8; perf_1** đến **perf_5; scor_1** đến **scor_5; attr_1 , attr_2, attr_3**.

Bảng 1. Các biến quan sát trong tệp dữ liệu *Corporate reputation data*.

Tên biến	Nội dung	Thang đo
	Năng lực (COMP)	
<i>comp_1</i>	[Công ty] là một đối thủ cạnh tranh hàng đầu trên thị trường.	1-7
<i>comp_2</i>	Theo tôi biết, [công ty] nổi tiếng toàn cầu.	1-7
<i>comp_3</i>	Tôi tin rằng [công ty] hoạt động ở mức cao cấp.	1-7
	Độ yêu thích (LIKE)	
<i>like_1</i>	[Công ty] là một công ty mà tôi có thể nhận dạng tốt hơn so với các công ty khác	1-7
<i>like_2</i>	[Công ty] là một công ty mà tôi sẽ hối tiếc nhiều nếu nó không còn tồn tại nữa vì tôi thích nó hơn các công ty khác.	1-7
<i>like_3</i>	Tôi xem [công ty] như là một công ty đáng để yêu thích.	1-7
	Sự hài lòng của khách hàng (CUSA)	
<i>cusa</i>	Nếu bạn xem xét trải nghiệm của bạn với [công ty], mức độ hài lòng của bạn đối với [công ty] như thế nào?	1-7
	Lòng trung thành của khách hàng (CUSL)	
<i>cusl_1</i>	Tôi sẽ giới thiệu [công ty] đến bạn bè và người thân của tôi.	1-7
<i>cusl_2</i>	Nếu phải lựa chọn lại, tôi cũng sẽ chọn [công ty] như nhà cung cấp dịch vụ điện thoại di động của tôi.	1-7
<i>cusl_3</i>	Tôi sẽ vẫn là một khách hàng [của công ty] trong tương lai.	1-7
	Chất lượng (QUAL)	
<i>qual_1</i>	Sản phẩm/dịch vụ được cung cấp bởi [công ty] có chất lượng cao.	1-7
<i>qual_2</i>	[Công ty] là một nhà đối mới chứ không phải bắt chước với sự tôn trọng [ngành].	1-7
<i>qual_3</i>	Sản phẩm/dịch vụ [của công ty] có giá tốt/phải chăng.	1-7
<i>qual_4</i>	Dịch vụ [công ty] cung cấp tốt.	1-7
<i>qual_5</i>	Các quan tâm của khách hàng được đánh giá cao tại [công ty].	1-7
<i>qual_6</i>	[Công ty] là đối tác đáng tin cậy đối với khách hàng.	1-7
<i>qual_7</i>	[Công ty] là một công ty đáng tin cậy.	1-7
<i>qual_8</i>	Tôi rất tôn trọng đối với [công ty].	1-7

	Sự hiệu quả (PERF)	
<i>perf_1</i>	[Công ty] là một công ty được quản trị rất tốt.	1-7
<i>perf_2</i>	[Công ty] là một công ty có kinh tế ổn định.	1-7
<i>perf_3</i>	Rủi ro kinh doanh của [công ty] là khiêm tốn nhất so với các đối thủ cạnh tranh.	1-7
<i>perf_4</i>	[Công ty] có tiềm năng tăng trưởng.	1-7
<i>perf_5</i>	[Công ty] có tầm nhìn rõ ràng về tương lai của công ty.	1-7
	Trách nhiệm xã hội của công ty (CSOR)	
<i>csor_1</i>	[Công ty] cư xử theo cách có ý thức về mặt xã hội.	1-7
<i>csor_2</i>	[Công ty] thăng thắn trong việc cung cấp thông tin cho công chúng.	1-7
<i>csor_3</i>	[Công ty] có thái độ công bằng đối với các đối thủ cạnh tranh.	1-7
<i>csor_4</i>	[Công ty] quan tâm đến việc bảo vệ môi trường.	1-7
<i>csor_5</i>	[Công ty] không chỉ quan tâm về lợi nhuận.	1-7
	Sự thu hút (ATTR)	
<i>attr_1</i>	[Công ty] thành công trong việc thu hút nhân sự chất lượng cao.	1-7
<i>attr_2</i>	Tôi có thể thấy mình đang làm việc tại [công ty].	1-7
<i>attr_3</i>	Tôi thích dáng vẻ bên ngoài của [công ty] (công ty, tòa nhà, cửa hàng, v.v.).	1-7
	Chi phí chuyển đổi (SWITCH)	
<i>switch_1</i>	Tôi mất rất nhiều thời gian để chuyển qua một công ty khác.	1-5
<i>switch_2</i>	Tôi tốn kém rất nhiều để chuyển qua một công ty khác.	1-5
<i>switch_3</i>	Tôi bỏ ra nhiều nỗ lực để làm quen với một công ty mới với các quy định và các cách làm việc của họ.	1-5
<i>switch_4</i>	Nói chung, sẽ là một sự phiền hà đáng kể để chuyển qua một công ty khác.	1-5

Nguồn: Hair và công sự. (2022: trang 70, 160-161)

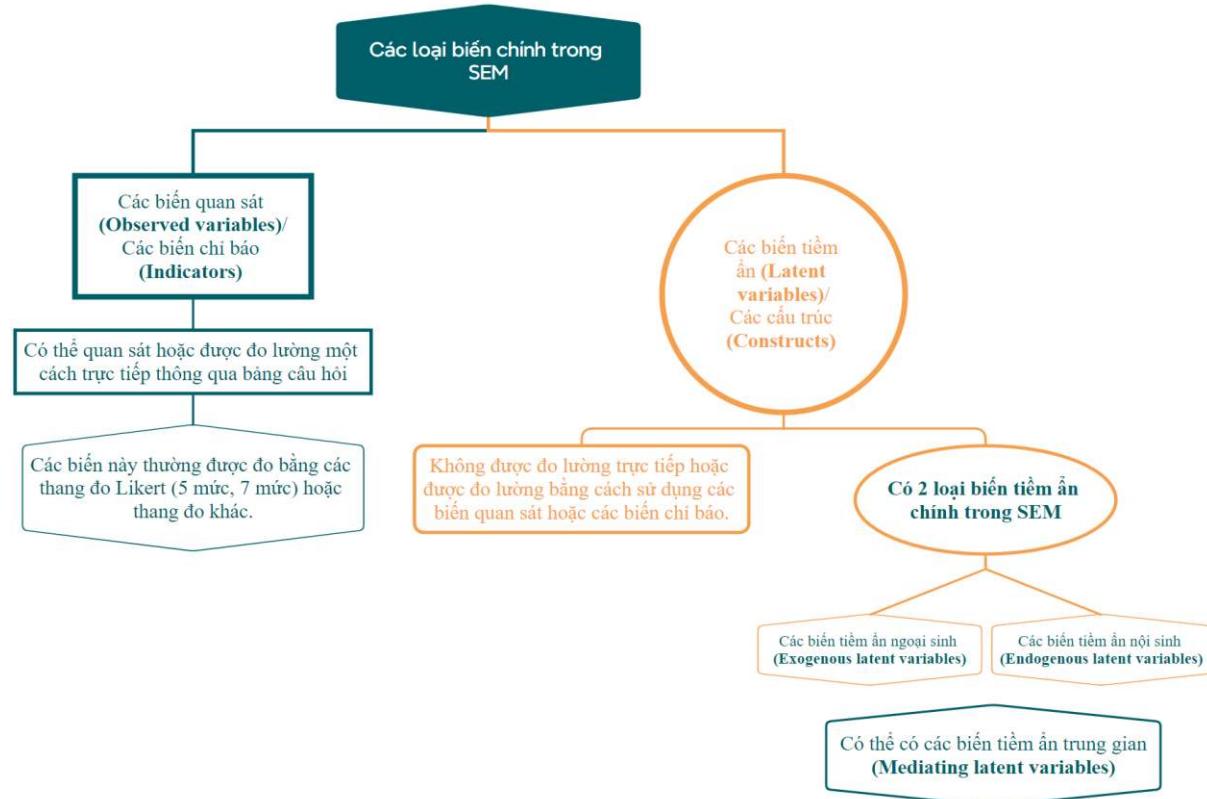
Ngoài các thang đo này, tệp dữ liệu còn có **biến quan sát chung** (global single-item measures) chỉ có một thang đo duy nhất để đánh giá tổng quát bốn khái niệm về chất lượng sản phẩm và dịch vụ của công ty (**qual_global**), hiệu quả kinh tế và quản trị của công ty (**perf_global**), trách nhiệm xã hội của công ty (**scor_global**), và sự thu hút của công ty (**attr_global**). Điều này có nghĩa là, trong bảng câu hỏi khảo sát, sau khi hỏi

các câu hỏi riêng biệt về chất lượng sản phẩm và dịch vụ của công ty (*qual_*) sẽ hỏi thêm một câu hỏi để đánh giá chung chất lượng sản phẩm và dịch vụ của công ty (*qual_global*); sau khi hỏi các câu hỏi riêng biệt về hiệu quả kinh tế và quản trị của công ty (*perf_*) sẽ hỏi thêm một câu hỏi để đánh giá chung hiệu quả kinh tế và quản trị của công ty (*perf_global*); sau khi hỏi các câu hỏi riêng biệt về trách nhiệm xã hội của công ty (*scor_*) sẽ hỏi thêm một câu hỏi để đánh giá chung trách nhiệm xã hội của công ty (*scor_global*); v.v. Thu thập thêm các biến quan sát chung này nhằm mục đích đo lường **sự hội tụ (convergent validity)** của các thang đo trong **mô hình đo lường nguyên nhân** bằng kỹ thuật phân tích dư thừa (**Redundancy analysis**). Các khái niệm này sẽ được giải thích và minh họa ở phần sau.

Ngoài ra, tệp dữ liệu còn có bốn biến quan sát để đo lường khái niệm nhận thức về **chi phí chuyển đổi (Switching costs)** từ việc sử dụng sản phẩm và dịch vụ của một công ty sang này sang một công ty khác. Biến tiềm ẩn/cấu trúc này sẽ được dùng trong Chapter 7 để minh họa về vai trò của **biến điều tiết (Moderator variable)** trong mối quan hệ giữa hai cấu trúc về sự hài lòng của khách hàng (CUSA) và lòng trung thành của khách hàng (CUSL). Các biến quan sát này được đo bằng thang đo Likert 5 điểm; được trình bày trong **Exhibit A7.2 (Case Study, Chapter 7: trang 4)** hoặc **Exhibit 7.22 (Hair và cộng sự., 2022: trang 261)**. Cuối cùng, tệp dữ liệu có biến giả về loại dịch vụ (hợp đồng hay trả trước), và nhà cung cấp dịch vụ điện thoại (5 nhóm).

3. GIỚI THIỆU MÔ HÌNH PLS-SEM

3.1 Các loại biến chính trong mô hình SEM



Hình 1. Các loại biến chính trong mô hình SEM, tổng hợp từ ([Ramlall, 2016: trang 4](#)).

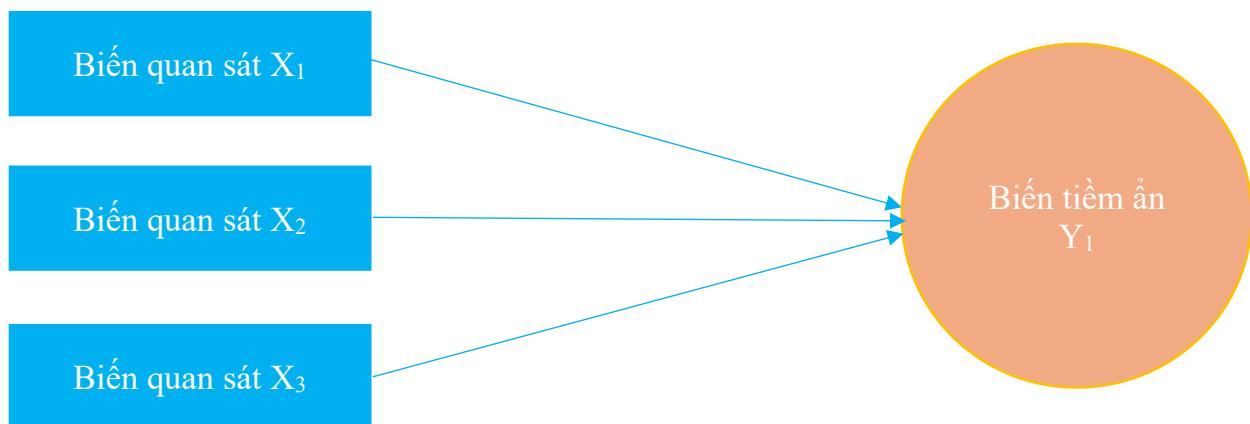
Hình 1 trình bày các loại biến chính trong mô hình SEM. *Thứ nhất*, các biến quan sát (**observed variables**) hay các biến chỉ báo (**indicators**) được thu thập trực tiếp từ bảng câu hỏi. Thông thường, các biến này được đo bằng thang đo Likert. *Thứ hai*, các biến tiềm ẩn (**latent variables**) hay các cấu trúc (**constructs**). Các biến này không thể đo lường hoặc thu thập một cách trực tiếp mà chỉ có thể được đo lường gián tiếp thông qua một hoặc nhiều biến quan sát. Nếu cấu trúc được đo bằng một biến quan sát duy nhất thì được gọi là đơn thang đo (**single-item**), và bằng nhiều biến quan sát thì được gọi là đa/nhiều thang đo (**multiple items**).

3.2 Các ký hiệu trong mô hình SEM

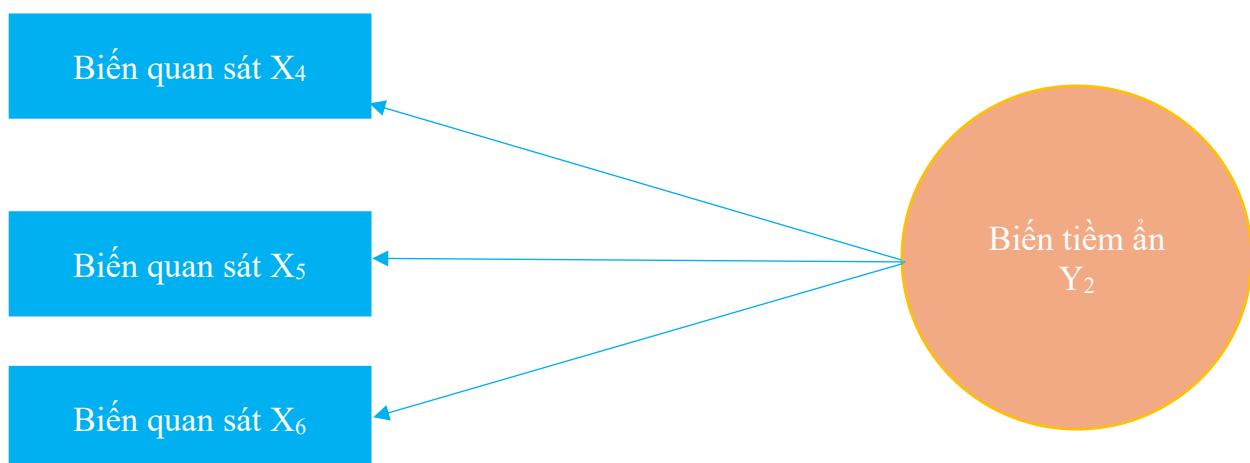
Theo Ramlall (2016: trang 30), các biến quan sát được biểu diễn bằng các hình chữ nhật, hình hộp hoặc hình vuông; các biến tiềm ẩn được biểu diễn bằng các hình Oval, hình Eclipse hoặc hình tròn.



Hình 2. Ký hiệu biến quan sát và biến tiềm ẩn.

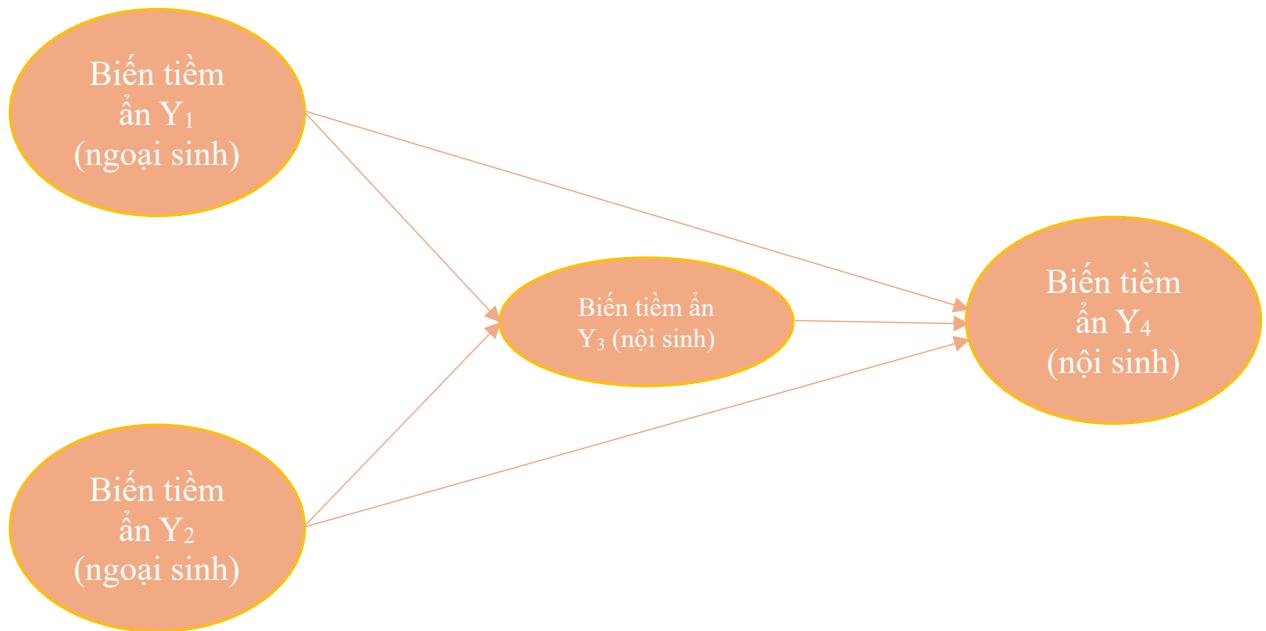


Hình 3. Mô hình đo lường nguyên nhân (Formative measurement model).



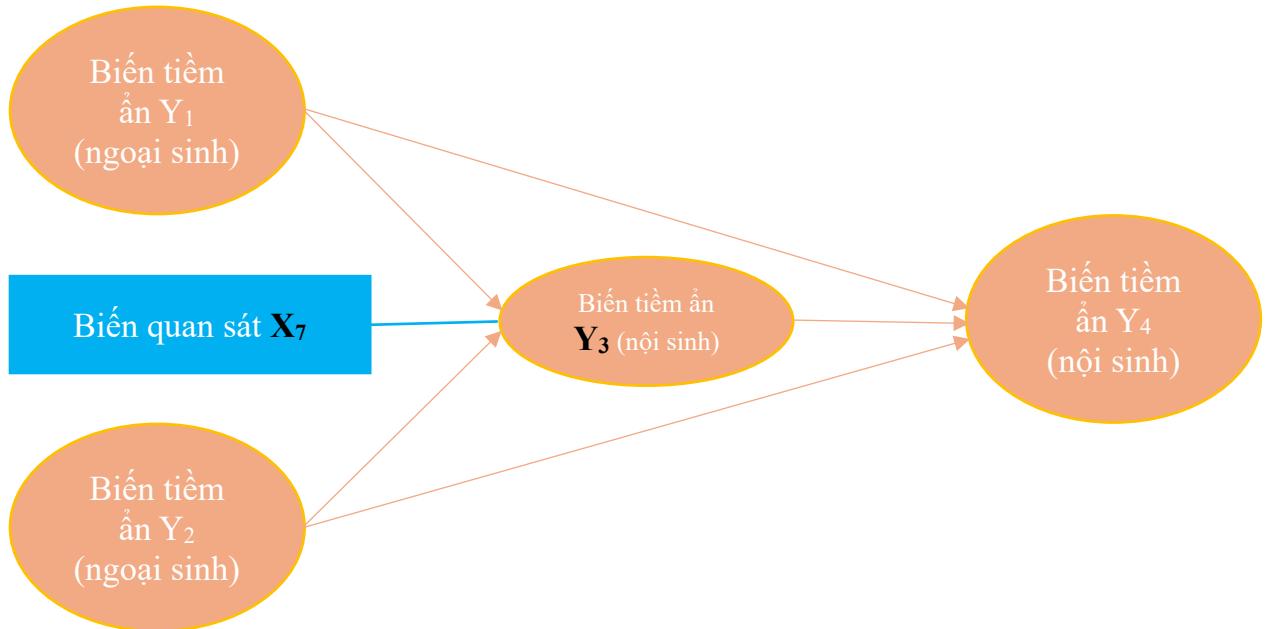
Hình 4. Mô hình đo lường kết quả (Reflective measurement model).

Mũi tên một chiều thể hiện mối quan hệ nguyên nhân/kết quả với nhiều hơn một biến quan sát (trong mô hình đo lường nguyên nhân: **Formative measurement model** hoặc mô hình đo lường kết quả: **Reflective measurement model**); hoặc đường dẫn (Path) từ một biến tiềm ẩn ngoại sinh (**Exogenous latent variable**) đến một biến tiềm ẩn nội sinh (**Endogenous latent variable**).



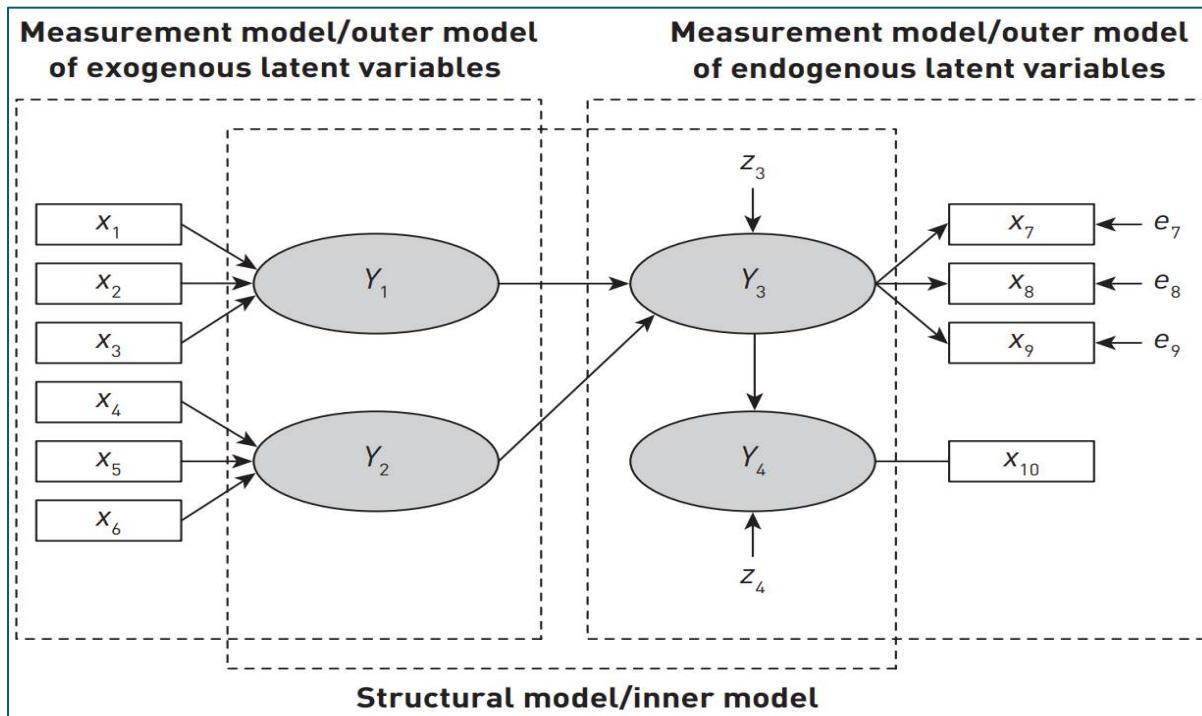
Hình 5. Mô hình cấu trúc (Structural model).

Nếu biến tiềm ẩn (như Y_3) chỉ gồm một biến quan sát duy nhất (X_7) thì mối quan hệ sẽ được biểu diễn bằng một đường gạch nối. Trường hợp này ít gặp trên thực tế. Tuy nhiên, đôi khi được sử dụng dưới dạng một **biến quan sát dạng tổng quát** (chung) ngay sau khi hỏi các câu hỏi riêng lẻ về các khía cạnh khác nhau của một khái niệm.



Hình 6. Biến tiềm ẩn (Y_3) chỉ gồm một biến quan sát duy nhất (X_7).

Ngoài ra, đường mũi tên còn được sử dụng để biểu diễn các **sai số** (error terms) trong các mô hình đo lường (như e_7 , e_8 , và e_9) và/hoặc mô hình cấu trúc (như z_3 , z_4) trong minh họa dưới đây (**Hình 7**).



Hình 7. Một mô hình PLS-SEM đơn giản (Hair và cộng sự., 2022: trang 12).

Theo Hair và cộng sự (2022: trang 12), trong mô hình **PLS-SEM**² các đường mũi tên luôn luôn là đơn đầu (**single-headed arrow**), và vì thế chúng thể hiện các mối quan hệ chiều hướng (**directional relationships**); chính xác hơn là các mối quan hệ dự đoán (**predictive relationships**). Với sự ủng hộ về mặt lý thuyết thì chúng giải thích các mối quan hệ nhân quả (**causal relationships**).

3.3 Tổng quan về mô hình PLS-SEM

Theo Hair và cộng sự (2022: trang 13), một mô hình đường dẫn PLS gồm hai thành phần (như **Hình 7**). *Thứ nhất*, đó là một mô hình cấu trúc (**structural model**) hay còn gọi là mô hình bên trong (**inner model**). Mô hình này kết nối các cấu trúc (các biến tiềm ẩn) với nhau (các hình tròn hoặc hình oval). Mô hình cấu trúc cũng biểu diễn các mối quan hệ (còn gọi là các đường dẫn - paths) giữa các cấu trúc (các biến tiềm ẩn). *Thứ hai*, một mô hình đo lường cấu trúc (**construct's measurement model**) hay còn gọi là mô hình bên ngoài (**outer model**). Mô hình này biểu diễn các mối quan hệ giữa các cấu trúc (các biến tiềm ẩn) với các biến chỉ báo (các biến quan sát). Như ở trên đã trình bày, các biến chỉ báo được thể hiện bằng các hình chữ nhật.

Như ở **Hình 7**, có hai loại **mô hình đo lường**: một cho **các biến tiềm ẩn ngoại sinh** (nghĩa là các cấu trúc này giải thích cho các cấu trúc khác trong mô hình), và một cho **các biến tiềm ẩn nội sinh** (nghĩa là các cấu trúc này được giải thích trong mô hình).

² Partial Least Squares Structural Equation Modeling = **Mô hình phương trình cấu trúc bình phương tối thiểu riêng phần** (Hair và cộng sự, 2018).

Ngoài ra, một cách phân loại phổ biến khác là: **mô hình đo lường kết quả** (**Reflective measurement model**) với hướng mũi tên từ một biến tiềm ẩn đến một biến quan sát (như $Y_3 \rightarrow X_7$) và có bao gồm **các sai số** (như e_7, e_8, e_9); và **mô hình đo lường nguyên nhân** (**Formative measurement model**) với mũi tên từ một biến quan sát đến một biến tiềm ẩn (như $X_1 \rightarrow Y_1$) và không bao gồm các sai số. Trong mô hình cấu trúc cũng bao gồm các sai số như z_3 và z_4 liên quan đến các biến tiềm ẩn nội sinh (Y_3 và Y_4).

Các biến quan sát/chỉ báo còn có các tên gọi khác như **manifest variables** (các biến quan sát) hoặc **items** (các biến thang đo) như biểu diễn trong **Hình 8**. Mỗi quan hệ giữa các biến trong mô hình PLS-SEM như trên **Hình 8** được biểu diễn dưới dạng phương trình như sau (xem [Mehmetoglu & Venturini, 2021: trang 100-101](#)):

$$Y_{3i} = p_0 + p_{13}Y_{1i} + p_{23}Y_{2i} + \zeta_i \quad (1)$$

$$Y_{2i} = w_{24}X_{4i} + w_{25}X_{5i} + w_{26}X_{6i} + \omega_i \quad (2)$$

$$X_{1i} = l_{11}Y_{1i} + \varepsilon_{i1} \quad (3)$$

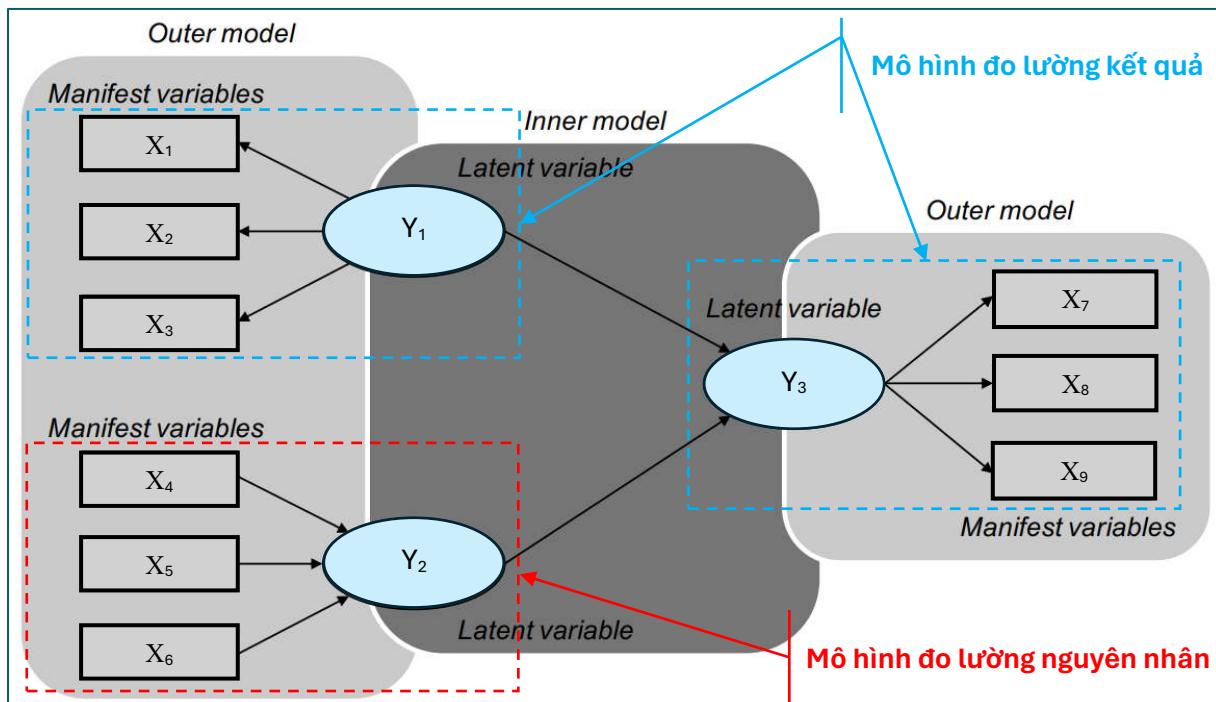
$$X_{2i} = l_{12}Y_{1i} + \varepsilon_{i2} \quad (4)$$

$$X_{3i} = l_{13}Y_{1i} + \varepsilon_{i3} \quad (5)$$

$$X_{7i} = l_{37}Y_{3i} + \varepsilon_{i7} \quad (6)$$

$$X_{8i} = l_{38}Y_{3i} + \varepsilon_{i8} \quad (7)$$

$$X_{9i} = l_{39}Y_{3i} + \varepsilon_{i9} \quad (8)$$

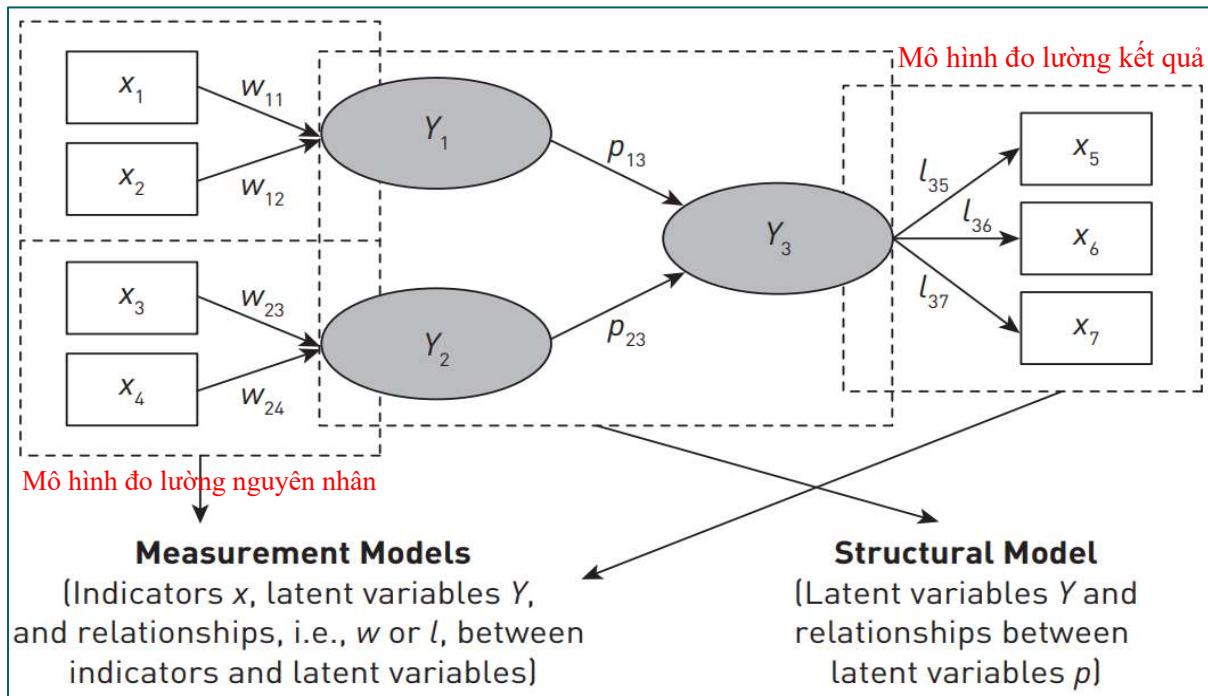


Hình 8. Một mô hình PLS-SEM đơn giản ([Mehmetoglu & Venturini, 2021: trang 94](#)).

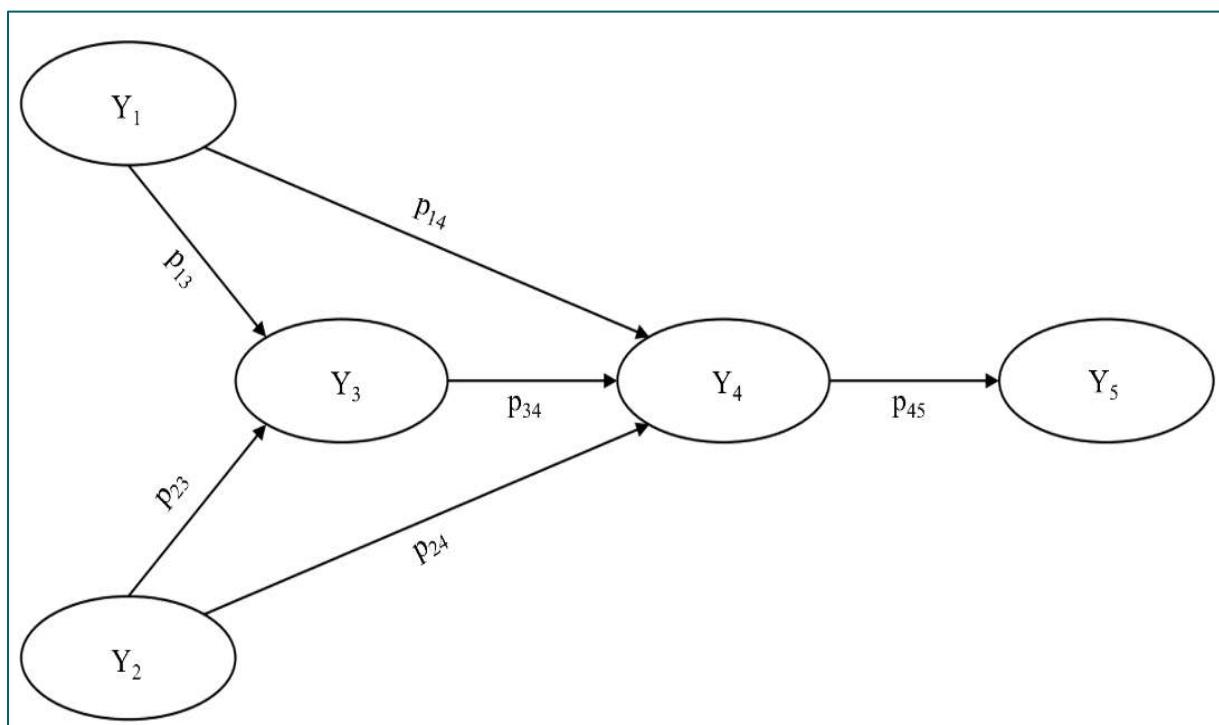
Các hệ số l được gọi là các **hệ số tải ngoài** (**outer loadings**) – thể hiện mức độ liên quan của một biến quan sát đến biến tiềm ẩn trong mô hình đo lường kết quả. Các hệ số w được gọi là các **trọng số ngoài** (**outer weights**) – thể hiện sự đóng góp tương đối của

một biến quan sát trong biến tiềm ẩn trong mô hình đo lường nguyên nhân. Các hệ số p là **hệ số đường dẫn** (path coefficients) – thể hiện **anh hưởng trực tiếp** (direct effect) của một biến tiềm ẩn ngoại sinh lên biến tiềm ẩn nội sinh trong mô hình cấu trúc.

Hair và cộng sự (2022: trang 87) biểu diễn các hệ số l , p , w trên mô hình PLS-SEM như hình dưới đây:



Hình 9. Biểu diễn các hệ số l , p , w trên mô hình PLS-SEM (Hair và cộng sự, 2022).



Hình 10. Một mô hình PLS-SEM phức tạp (Mehmetoglu & Venturini, 2021: trang 125).

Với mô hình PLS-SEM phức tạp hơn, trong có một hay nhiều biến tiềm ẩn đóng vai trò vừa là biến giải thích vừa là biến phụ thuộc (như Y₃ và Y₄ ở **Hình 10**) – biến trung gian, thì sẽ có ba phương trình của mô hình cấu trúc tương ứng với ba biến tiềm ẩn nội sinh (Y₃, Y₄, và Y₅). Các phương trình này như sau (xem thêm Muller và cộng sự, 2005):

$$\text{Y}_{3i} = p_0 + p_{13}Y_{1i} + p_{23}Y_{2i} + \zeta_{i3} \quad (9)$$

$$\text{Y}_{4i} = p_0 + p_{14}Y_{1i} + p_{24}Y_{2i} + p_{34}\text{Y}_{3i} + \zeta_{i4} \quad (10)$$

$$\text{Y}_{5i} = p_0 + p_{45}\text{Y}_{4i} + \zeta_{i5} \quad (11)$$

Để đơn giản, ta chỉ tập trung vào hai mối quan hệ: ảnh hưởng của Y₁ lên Y₄ và ảnh hưởng của Y₂ lên Y₅.

- Ảnh hưởng trực tiếp của Y₁ lên Y₄ là hệ số đường dẫn p₁₄.
- **Ảnh hưởng gián tiếp** (indirect effect) của Y₁ lên Y₄ được thể hiện bằng ảnh hưởng của Y₁ lên Y₄ thông qua Y₃; được đo bằng tích của hai hệ số đường dẫn tương ứng, đó là (p₁₃ x p₃₄).
- **Tổng ảnh hưởng** (total effect) của Y₁ lên Y₄ là p₁₄ + (p₁₃ x p₃₄).
- Không có ảnh hưởng trực tiếp của Y₂ lên Y₅ bởi vì không có đường dẫn trực tiếp từ Y₂ đến Y₅.
- Ảnh hưởng gián tiếp của Y₂ lên Y₅ thông qua hai đường dẫn: (i) từ Y₂ qua Y₅ thông qua Y₄ (nghĩa là Y₂ → Y₄ → Y₅), và hệ số ảnh hưởng gián tiếp này là (p₂₄ x p₄₅); và (ii) từ Y₂ qua Y₅ thông qua Y₃ và Y₄ (nghĩa là Y₂ → Y₃ → Y₄ → Y₅), và hệ số ảnh hưởng gián tiếp này là (p₂₃ x p₃₄ x p₄₅). Vì thế, **tổng ảnh hưởng gián tiếp** (overall indirect effect) của Y₂ lên Y₅ là tổng của hai ảnh hưởng gián tiếp: (p₂₄ x p₄₅) + (p₂₃ x p₃₄ x p₄₅).
- Do không có ảnh hưởng trực tiếp, nên tổng ảnh hưởng của Y₂ lên Y₅ cũng chính là tổng ảnh hưởng gián tiếp: (p₂₄ x p₄₅) + (p₂₃ x p₃₄ x p₄₅).

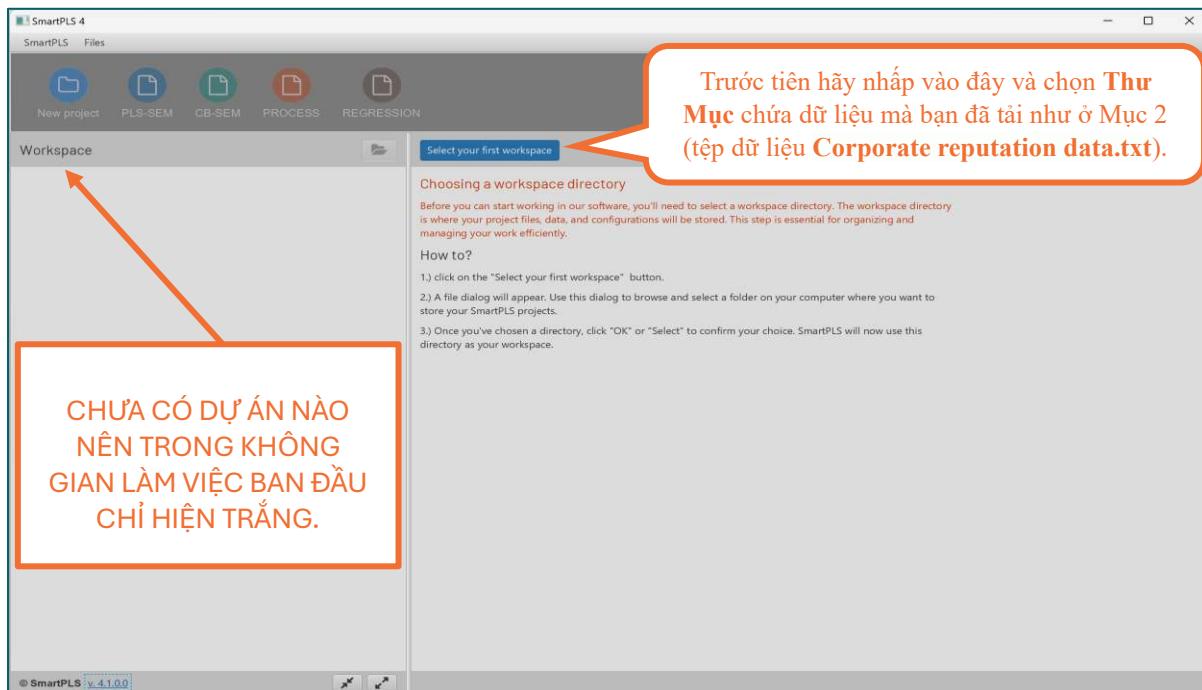
Trong **Hình 10** này, Y₃, Y₄, và Y₅ đều là các biến tiềm ẩn nội sinh, nhưng Y₃ và Y₄ cũng giữ vai trò là các biến tiềm ẩn ngoại sinh. Theo ngôn ngữ của SEM, thì Y₃ và Y₄ là các **biến trung gian** (mediating variables). Một lưu ý khác là các mô hình đo lường không có liên quan gì đến định nghĩa các ảnh hưởng trực tiếp và gián tiếp (Mehmetoglu & Venturini, 2021: trang 125). Việc xác định mô hình nghiên cứu có (các) biến trung gian hay không tùy vào cơ sở lý thuyết. Thông thường thì các khung lý thuyết về hành vi như Lý thuyết hành vi dự định (TPB), Lý thuyết hành động hợp lý (TRA), Lý thuyết mô hình chấp nhận công nghệ (TAM), Lý thuyết động cơ bảo vệ (PMT), Lý thuyết giá trị kỳ vọng (EVT) đều có các biến trung gian.

4. TẠO MÔ HÌNH PLS-SEM TRÊN SMARTPLS 4

4.1 Giao diện SmartPLS 4 và chuẩn bị dữ liệu

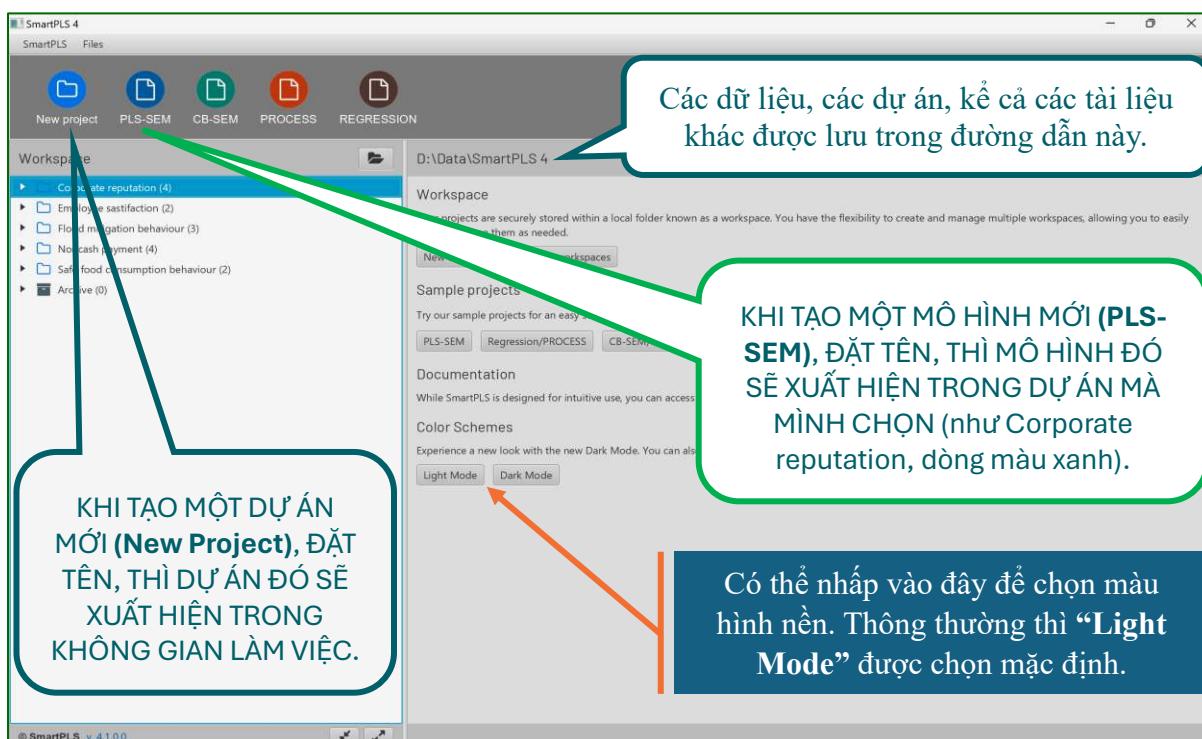
Hình 11 là giao diện của SmartPLS 4 sau khi vừa cài đặt. Việc đầu tiên bạn nên làm là nhấp vào “**Select your first workspace**” – một thư mục (folder) mà bạn tạo sẵn để chứa các dữ liệu cũng như các dự án nghiên cứu sau này. Ví dụ, tôi tạo một thư mục với tên **SmartPLS 4** chứa trong thư mục **Data** trong ô đĩa **D** (**D:\Data\SmartPLS 4**). Sau khi nhấp vào “**Select your first workspace**” và chọn đường dẫn như trên, thì giao diện

SmartPLS 4 sẽ hiện đường dẫn là D:\Data\SmartPLS 4 và phía dưới không gian làm việc (**Workspace**) bây giờ xuất hiện các dự án đã được thực hiện như **Hình 12**.



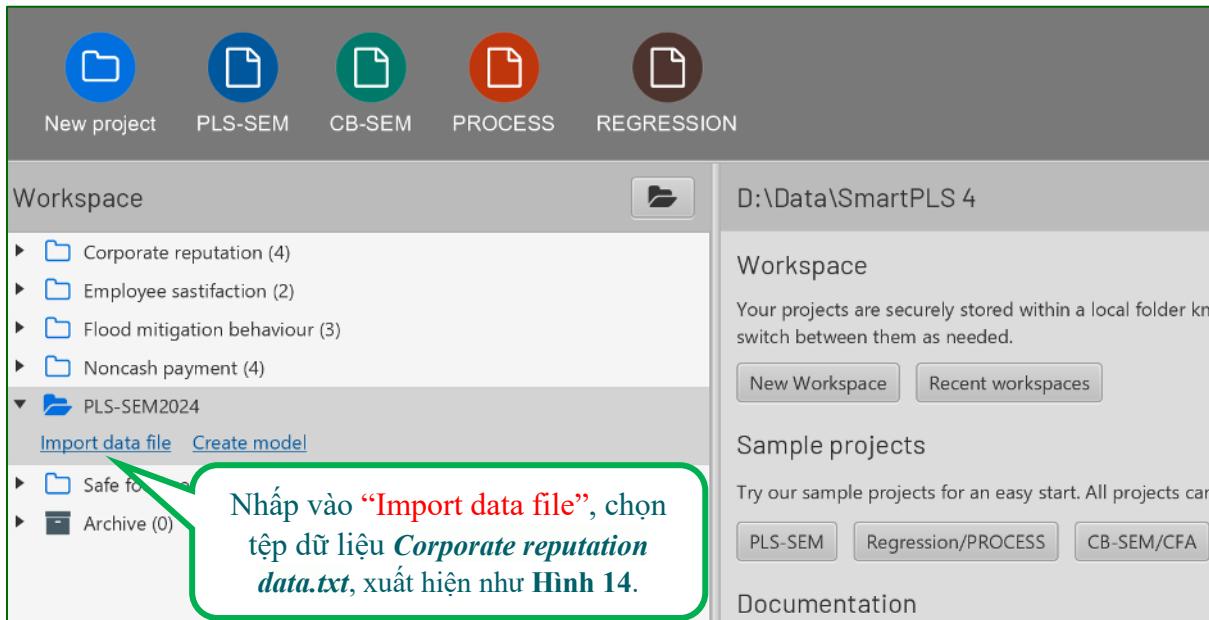
Hình 11. Giao diện SmartPLS 4 ban đầu.

Ban đầu thì không gian làm việc của bạn sẽ trống trơn như ở gốc trái **Hình 11**. Khi đã xác định thư mục cụ thể cho không gian làm việc thì các dự án sau này (kể cả dữ liệu và các kết quả ước lượng) sẽ tự động được lưu trong không gian làm việc đó và tiêu đề trên giao diện bên phải sẽ hiện ra đường dẫn các thư mục (như **Hình 12**).



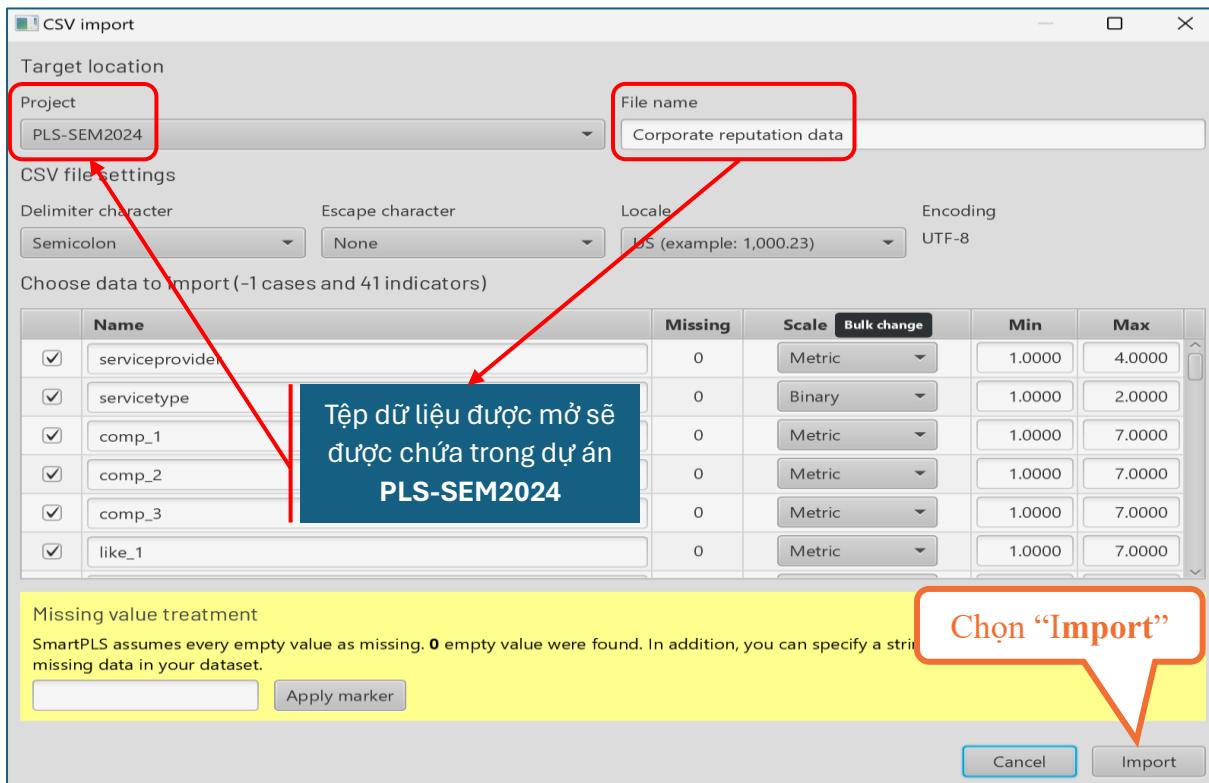
Hình 12. Giao diện SmartPLS 4 sau khi đã chọn thư mục D:\Data\SmartPLS 4.

Bây giờ, nhấp vào biểu tượng “**New Project**”, khi xuất hiện hộp thoại “**Create a new project**”, (tôi) đặt tên dự án này là **PLS-SEM2024** trong ô “**Name**”, rồi chọn “**Create**”. Kết quả xuất hiện như **Hình 13**.



Hình 13. Tạo một dự án mới trên SmartPLS 4.

Trường hợp bạn đã lưu tệp dữ liệu **Corporate reputation data.txt** trong không gian làm việc (**D:\Data\SmartPLS 4**), thì khi nhấp vào “**Import data file**” bạn sẽ được dẫn ngay đến đường dẫn đó để chọn tệp dữ liệu.



Hình 14. Nhập dữ liệu dạng csv hoặc xlsx vào SmartPLS 4.

Chọn “Back” để trở về không gian làm việc.

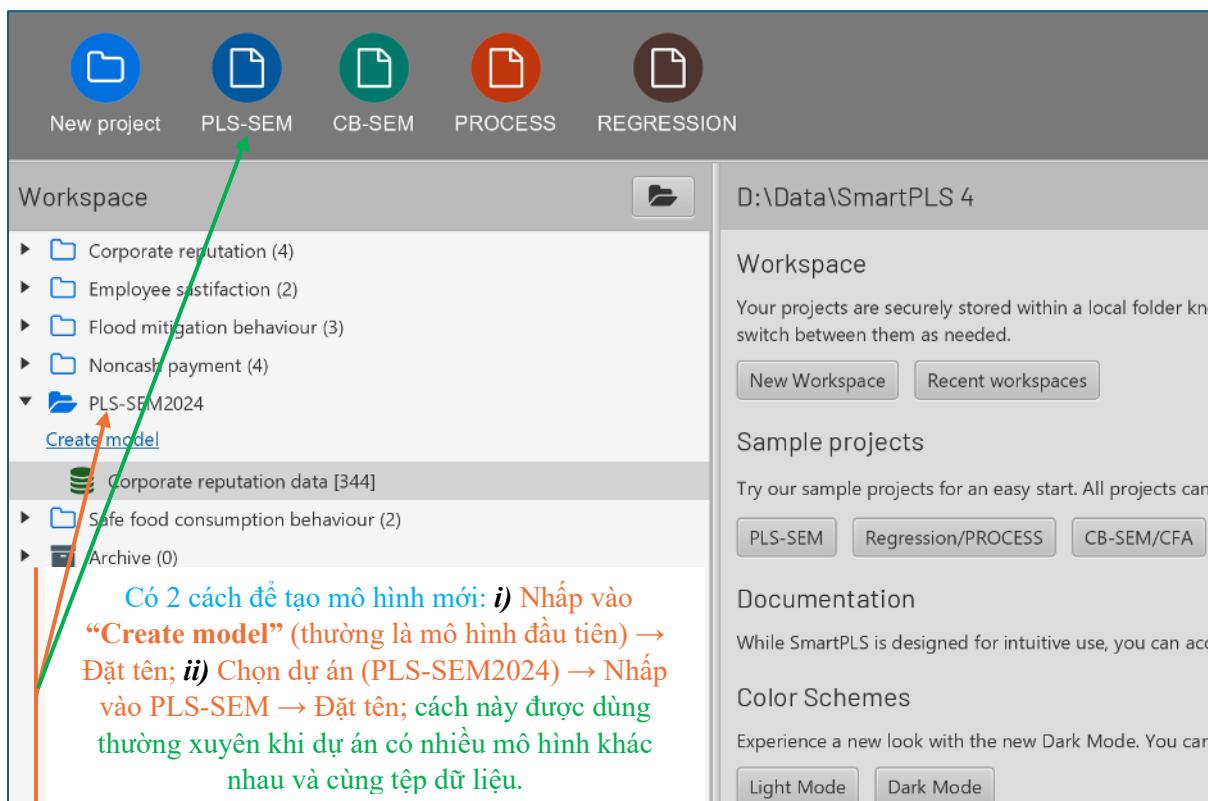
MET = Multi-Edge-Type, như thang đo Likert 5 điểm, 7 điểm, ...

Nếu chọn “Raw data” thì ta sẽ thấy file dữ liệu gốc, với dòng trên cùng là tên của các biến quan sát.

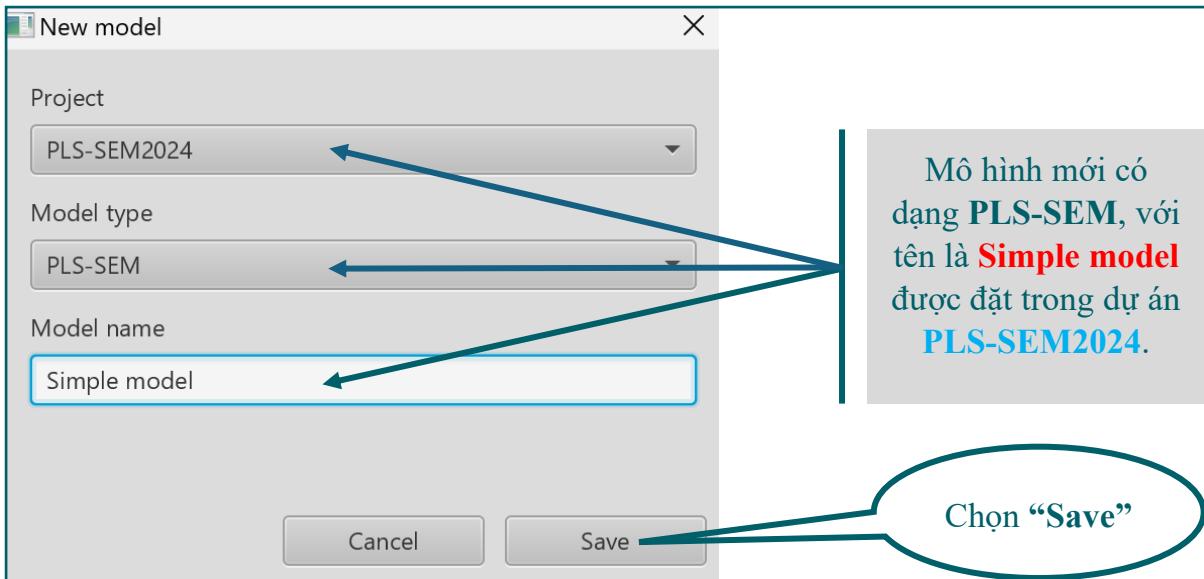
	Name	No.	Type	Missings	Mean	Median	Scale min	Scale max	Observed min	Observed max	Standard deviation
values	comp_3	1	MET	0	2.000	2.000	1.000	4.000	1.000	4.000	1.003
	like_1	2	0 1	0	1.637				0.00	2.000	0.481
	like_2	3	MET	0	5.42				0.00	7.000	1.433
	like_3	4	MET	0	5.42				0.00	7.000	1.375
Indicators	cusL1	5	MET	0	5.22				0.00	7.000	1.458
Correlations	cusL2	6	MET	0	4.584				0.00	7.000	1.547
Data groups	cusL3	7	MET	0	4.250	4.000	1.000	7.000	1.000	7.000	1.848
Raw data	cusa	8	MET	0	4.480	5.000	1.000	7.000	1.000	7.000	1.871
	csor_1	9	MET	0	4.221	5.000	-99.000	7.000	-99.000	7.000	9.798
	csor_2	10	MET	0	4.064	6.000	-99.000	7.000	-99.000	7.000	11.312
	csor_3	11							0.00	7.000	9.869
	csor_4	12							0.00	7.000	5.744
		13							0.00	7.000	1.469
		14							0.00	7.000	1.651
		15	MET	0	3.988	4.000	1.000	7.000	1.000	7.000	1.478
		16	MET	0	3.125	3.000	1.000	7.000	1.000	7.000	1.462

Hình 15. Bảng thống kê mô tả dữ liệu sau khi “import” dữ liệu.

4.2 Tạo một mô hình PLS-SEM mới

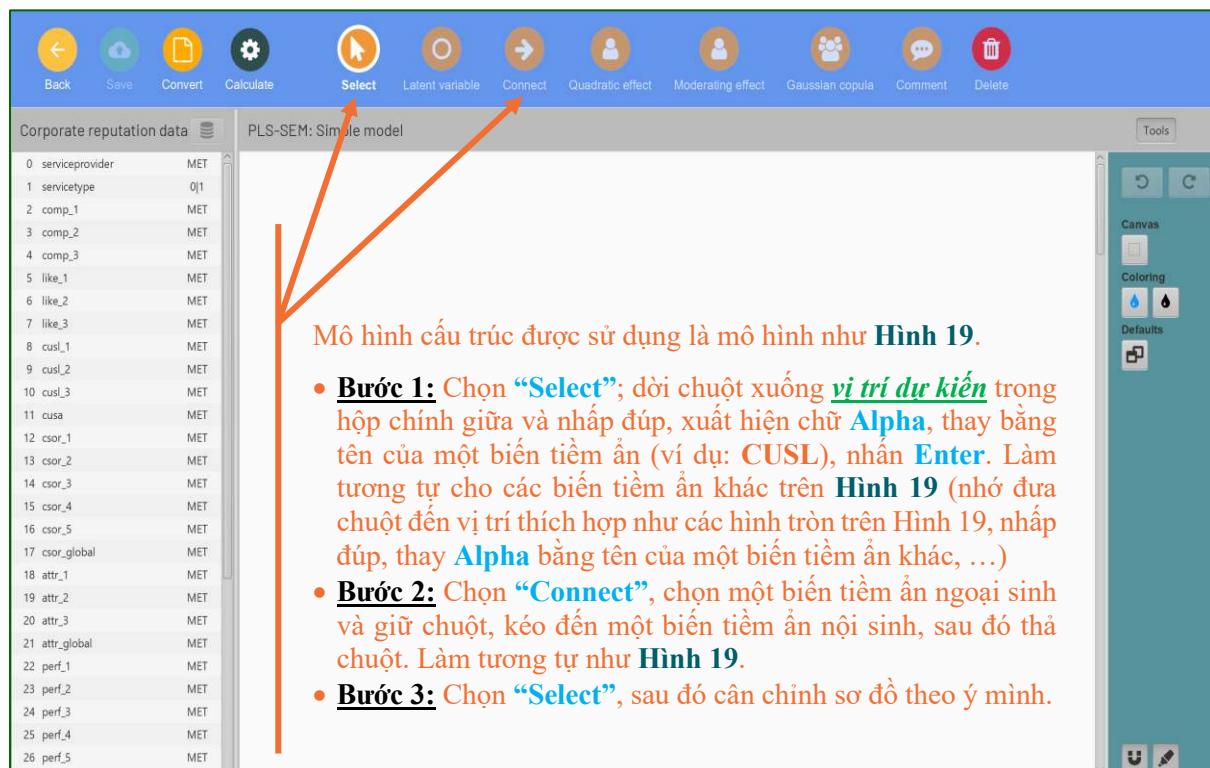


Hình 16. Tạo một mô hình PLS-SEM mới.



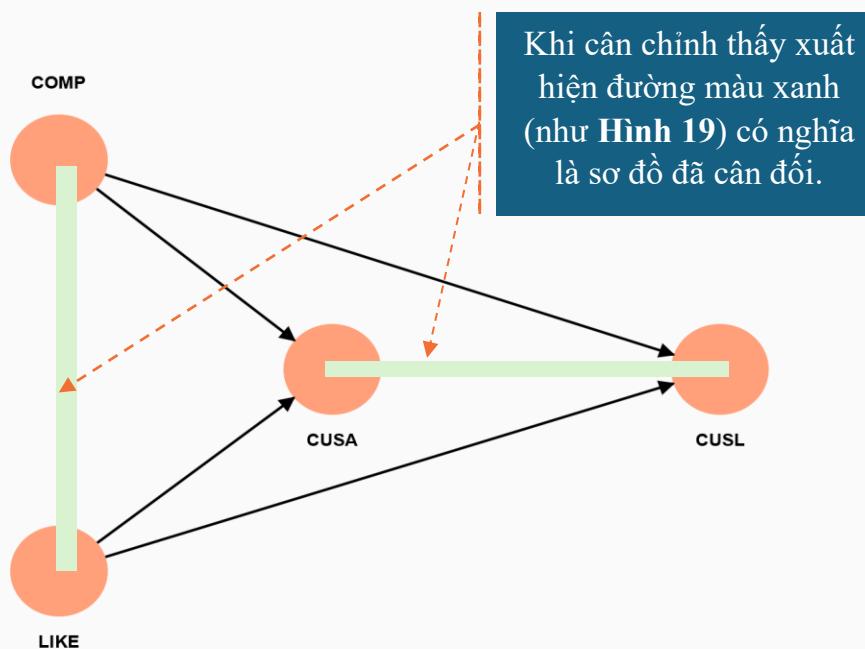
Hình 17. Hộp thoại mô hình PLS-SEM mới.

Sau khi chọn **“Save”** sẽ xuất hiện một giao diện làm việc mới như **Hình 18**, trong đó hộp bên phải là tất cả các biến quan sát (các biến chỉ báo) trong tệp dữ liệu, hộp chính giữa là nơi để vẽ sơ đồ của mô hình cấu trúc (kết nối các biến tiềm ẩn/cấu trúc với nhau, và sau đó là chọn các biến quan sát cho mỗi biến tiềm ẩn/cấu trúc), và bên phải là hộp công cụ cho việc định dạng.



Hình 18. Tạo sơ đồ mô hình cấu trúc theo mục tiêu nghiên cứu.

Mô hình cấu trúc ở **Hình 19** được trích từ **Exhibit 2.14** trong **Hair và cộng sự (2022: trang 68)** và **Exhibit A2.1** trong **Case Study, Chapter 2 (trang 3)**.

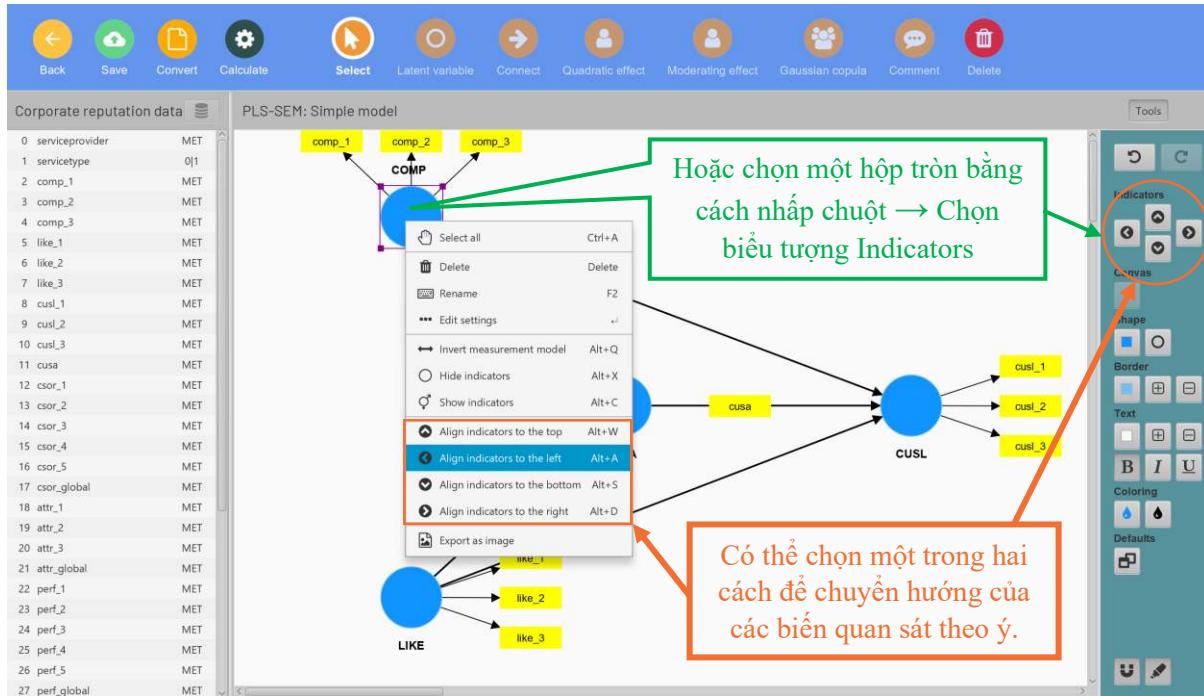


Hình 19. Tạo mô hình đơn giản về lòng trung thành Hair và cộng sự (2022: trang 68).

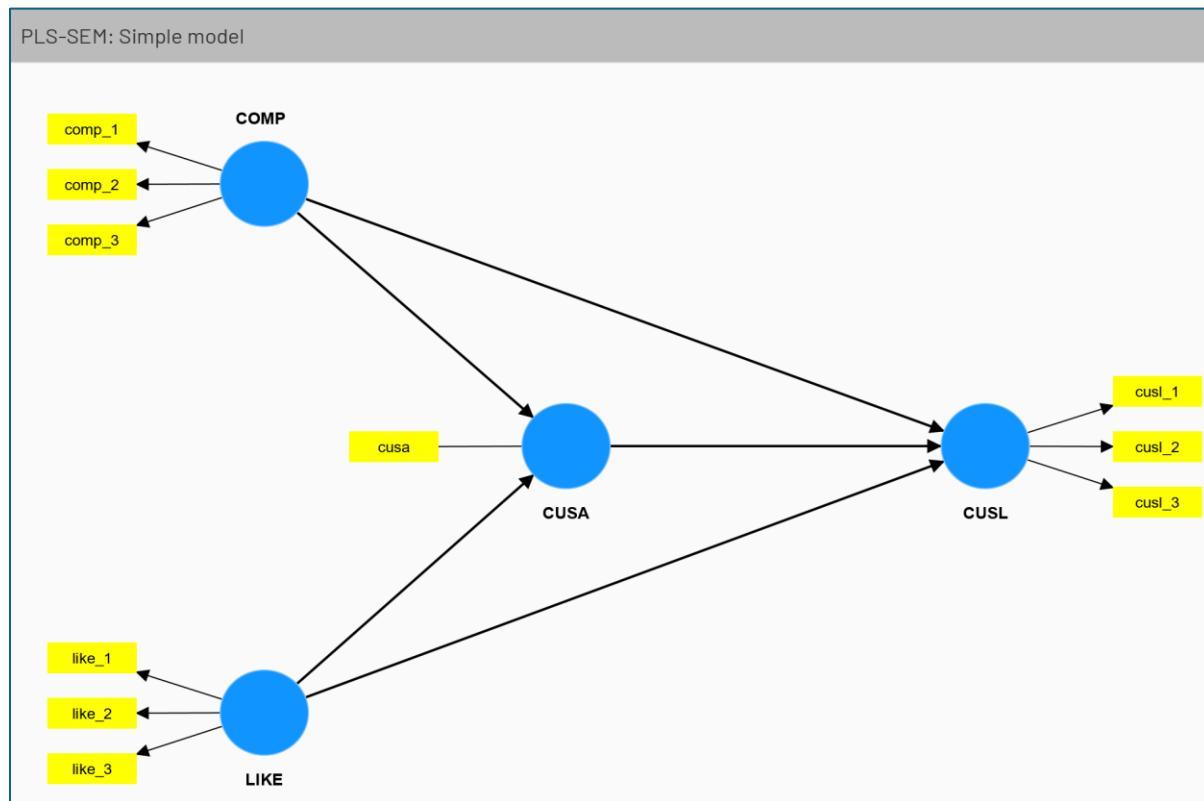
Bước tiếp theo là xác định mô hình đo lường tương ứng với các biến tiềm ẩn vừa được tạo ra như ở **Hình 19**. Cách thực hiện như sau:

- Chọn “**Select**”, sau đó nhấp chuột vào một biến tiềm ẩn bất kỳ trong sơ đồ mô hình cấu trúc (ví dụ: **COMP**).
- Chọn các biến quan sát tương ứng với biến tiềm ẩn **COMP** trong hộp bên trái (**comp_1 comp_2 comp_3**), giữ chuột và kéo bỏ vào hộp tròn của biến tiềm ẩn COMP. Khi này thì màu của hình tròn COMP sẽ chuyển sang màu xanh và các biến quan sát **comp_1 comp_2 comp_3** trong các hộp chữ nhật màu vàng. Nếu muốn bỏ màu hoặc đổi màu thì ta chọn “**Select**”, sau đó chọn các hình muốn đổi màu, hộp công cụ bên phải xuất hiện màu của **Shape**, **Border**, và **Text**, và ta có thể chọn màu khác. Làm tương tự cho các biến tiềm ẩn còn lại trong sơ đồ mô hình cấu trúc.
- Các biến quan sát khi được bỏ vào hộp tròn của biến tiềm ẩn tương ứng sẽ có dạng **mặc định** là **mô hình đo lường kết quả** (nghĩa là mũi tên đi từ biến tiềm ẩn đến biến quan sát). Nếu muốn chuyển sang mô hình đo lường nguyên nhân, ta đặt chuột vào hình tròn của một biến tiềm ẩn (ví dụ: **COMP**), nhấp phải, chọn “**↔ Invert measurement model**”. Tuy nhiên, trong mô hình đơn giản (**Simple model**) này ta chỉ thực hiện với mô hình đo lường kết quả nên chưa cần đến thao tác này. Mô hình mở rộng (**Extended model**) sẽ bao gồm cả mô hình đo lường kết quả và mô hình đo lường nguyên nhân.

- Kết quả như ở **Hình 20** là chưa được như ý. Ta cần sắp xếp lại các biến quan sát (ở đây gọi là **Indicators**). Ví dụ, ta muốn chuyển các biến quan sát của **COMP** sang trái, thì ta nhấp chuột vào hộp tròn **COMP**, sau đó *chọn mũi tên bên trái ở Indicators*. Làm tương tự ta có một sơ đồ hoàn chỉnh như **Hình 21**.



Hình 20. Tạo mô hình PLS-SEM đơn giản Hair và cộng sự (2022: trang 78).



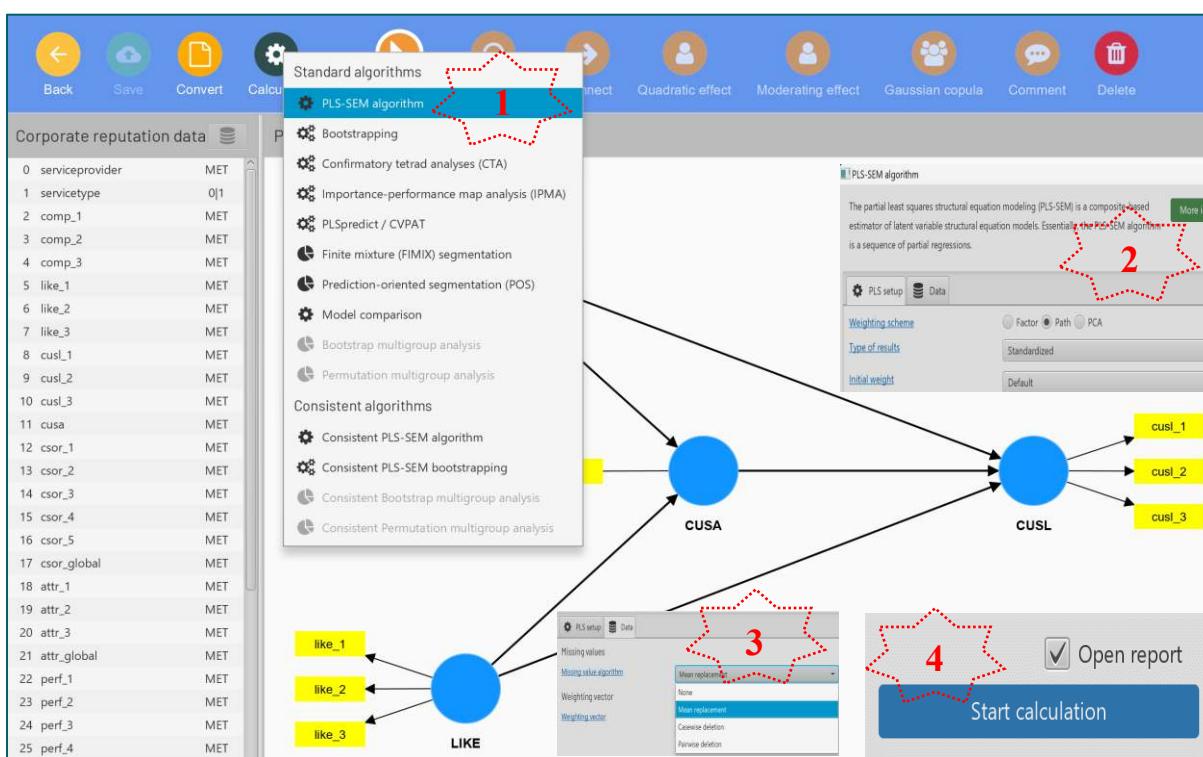
Hình 21. Tạo mô hình đơn giản về lòng trung thành Hair và cộng sự (2022: trang 78).

Sau cùng, ta có được một mô hình PLS-SEM đơn giản hoàn chỉnh ở **Hình 21** như trong **Exhibit 2.10** (Hair và cộng sự, 2022: trang 78) hoặc **Exhibit A2.8** (Case Study, Chapter 2: trang 13). Bây giờ, ta có thể sử dụng mô hình này cho việc ước lượng và sau đó là đánh giá các kết quả ước lượng. Tuy nhiên, hãy bấm “**SAVE**” để lưu mô hình lại. Và mặc định thì mô hình vừa được tạo ra sẽ được lưu trong **dự án PLS-SEM2024**, chưa trong không gian làm việc.

5. ƯỚC LƯỢNG MÔ HÌNH PLS-SEM TRÊN SMARTPLS 4

5.1 Ước lượng mô hình PLS-SEM

Sau khi đã hoàn thành việc tạo mô hình PLS-SEM, việc tiếp theo là tiến hành việc ước lượng. Có hai thuật toán để ước lượng mô hình PLS-SEM: Các thuật toán chuẩn (**Standard algorithms**) và các thuật toán nhất quán (**Consistent algorithms**). Phần này sẽ trình bày các thuật toán chuẩn. Với SmartPLS 4, ta sẽ nhấp vào “**Calculate**”, rồi chọn “**PLS-SEM algorithm**” như **Hình 22**.



Hình 22. Ước lượng mô hình đơn giản về lòng trung thành của khách hàng.

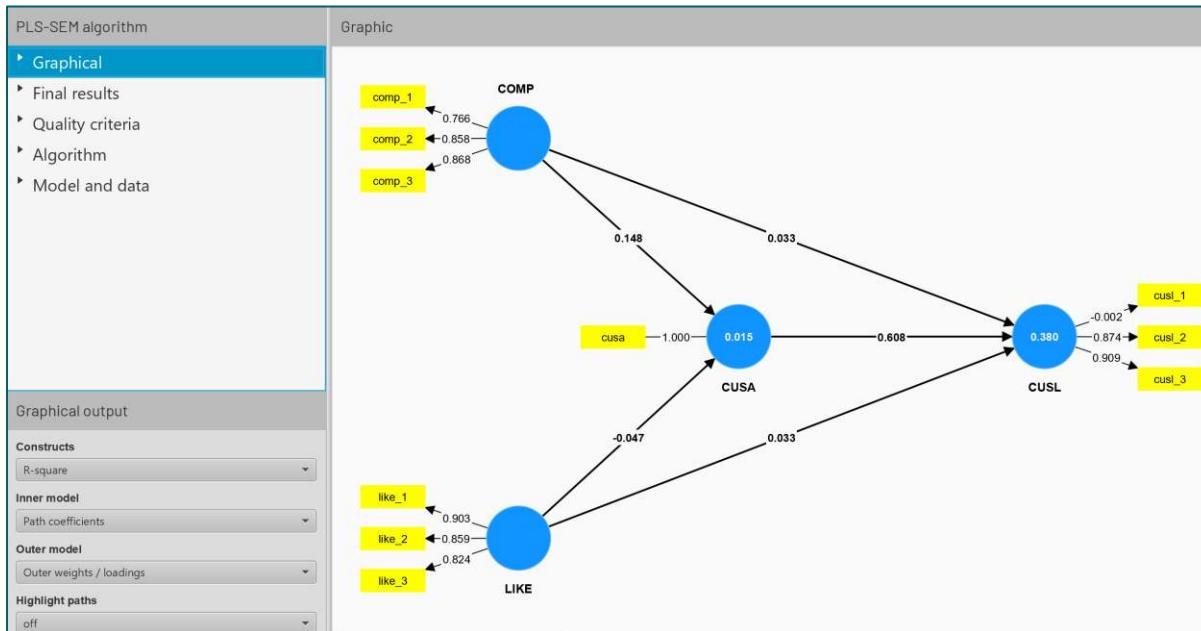
Trong “**PLS Setup**”, chọn “**Path**” trong “**Weighting schemes**”, “**Standardized**” trong “**Type of results**”³, và “**Default**” trong “**Initial weight**” một cách mặc định. Trong “**Data**”, nếu phần mềm phát hiện bộ dữ liệu có giá trị quan sát bị thiếu thì *tư đồng hiện thi* “**Missing values**” và mặc định là chọn “**Mean replacement**”. Theo Hair và cộng sự (2022: trang 95), thuật toán PLS-SEM dựa trên các điểm số của biến tiềm ẩn được chuẩn

³ Trong Case Study, Chapter 3: trang 2 có lưu ý rằng nếu chọn “**Unstandardized**” thì các hệ số không chuẩn hóa của các mô hình cấu trúc được giải thích theo sự thay đổi trong các giá trị của biến tiềm ẩn thay vì các độ lệch chuẩn. Điều này chỉ hợp lý khi tất cả các biến chỉ báo đều cùng một loại thang đo (như cùng thang Likert 7 điểm, hoặc cùng thang đo 5 điểm).

hóa (standardized). Vì thế, các ứng dụng của PLS-SEM phải sử dụng dữ liệu chuẩn hóa cho các biến chỉ báo (cụ thể hơn, đó là chuẩn hóa theo phân phối z, ở đó mỗi biến chỉ báo có giá trị trung bình bằng 0 và phương sai bằng 1) làm đầu vào cho thuật toán.

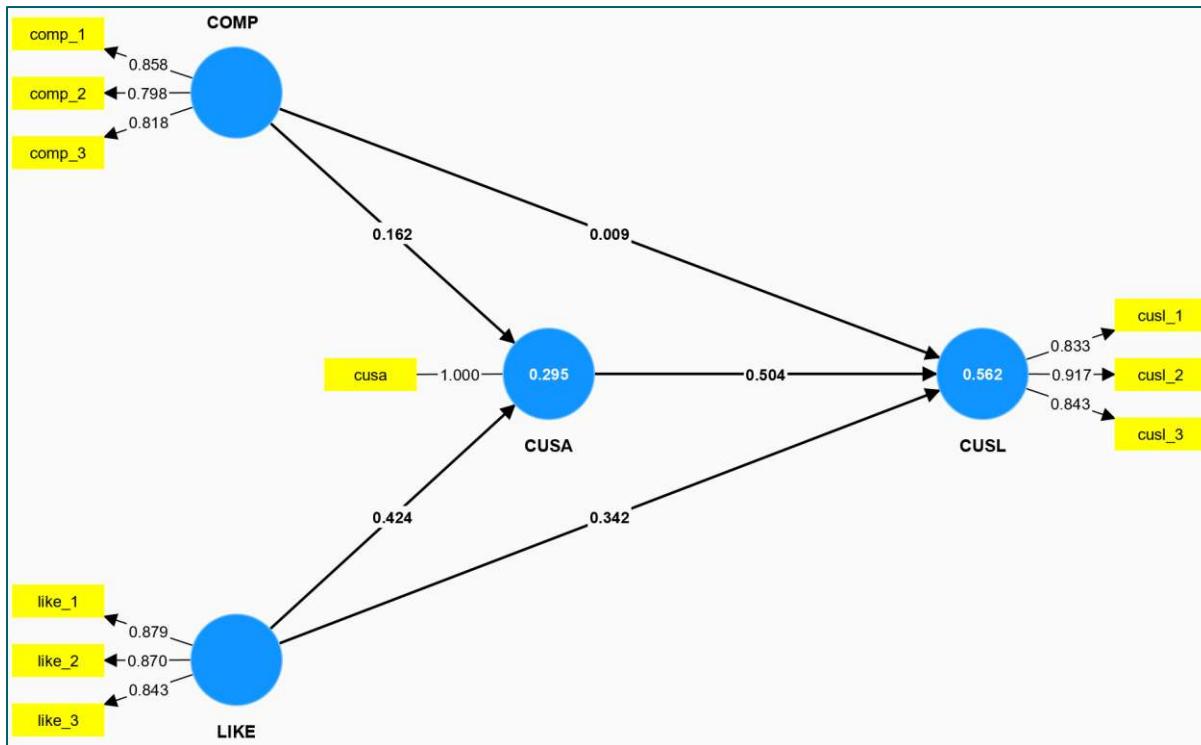
Kết quả là các hồi quy riêng phần được sử dụng trong các ước lượng mô hình có hệ số cắt (**Intercept**) bằng 0. Khi chạy phương pháp PLS-SEM thì phần mềm SmartPLS 4 chuẩn hóa cả dữ liệu của các biến chỉ số và các điểm số của biến tiềm ẩn. Điều này dẫn đến các hệ số ước lượng đều là các hệ số chuẩn hóa (**standardized coefficients**) trong mọi mối quan hệ trong mô hình cấu trúc (tức các hệ số đường dẫn, **p**) và các mô hình đo lường (tức các hệ số tải ngoài, **I** và các trọng số ngoài, **w**). Các hệ số này sẽ có giá trị trong khoảng -1 đến +1. Ví dụ, các hệ số đường dẫn gần bằng +1 cho biết một mối quan hệ dương mạnh giữa hai biến tiềm ẩn (ngược lại đối với các giá trị âm). Các hệ số càng gần 0 thì mối quan hệ giữa hai biến càng yếu. Các giá trị rất thấp gần bằng 0 thường là không có ý nghĩa thống kê. Để có các giá trị xác suất **p** (**p-values**) hoặc giá trị thống kê **t** (**t-stats**), thì ta phải ước lượng lại mô hình PLS-SEM bằng kỹ thuật mô phỏng (**Bootstrapping**). Ta sẽ đề cập đến vấn đề này ở phần sau.

Sau cùng là bấm “**Start calculation**”. Việc đầu tiên cần làm ngay sau khi hoàn thành thao tác ước lượng là lưu kết quả (“**SAVE**”, xuất hiện hộp thoại “**Save reports**”, đặt tên cho tệp kết quả ước lượng theo gợi ý mặc định: Tên mô hình + PLS + Năm–Tháng–Ngày, hoặc bất kỳ tên nào cũng được) trong không gian làm việc hiện hành. Kết quả xuất hiện đầu tiên là ở dạng đồ thị (**Graphical output**) như ở **Hình 23**.



Hình 23. Kết quả ước lượng mô hình đơn giản về lòng trung thành của khách hàng⁴.

⁴ Các kết quả ước lượng này có khác so với các kết quả ước lượng được trình bày trong Hair và cộng sự (2022: trang 102) và Case Study, Chapter 3: trang 7. **Tại sao?** Bởi vì trong bộ dữ liệu có một số “**missing values**” ở các biến quan sát *cusl_1*, *cusl_2*, *cusl_3*, và *cusa* được gán các giá trị **-99**. Với các giá trị **-99** thì SmartPLS 4 không xem là các giá trị bị thiếu (missing values), nên trong lựa chọn “**Data**” không xuất hiện “**Missing values**”. **Vậy ta phải làm gì?** Trong hướng dẫn này, tôi mở tệp dữ liệu “**Corporate reputation data**” bằng **Stata**, thay thế các giá trị **-99** bằng dấu “.”, sau đó lưu thành một tệp dữ liệu mới “**Corporate reputation data_new**”; Vào **Simple model** trong **PLS-SEM2024**, chọn “**Corporate reputation data**”, xóa; Chọn **Import data file** với tệp dữ liệu mới “**Corporate reputation data_new**”. Kết quả ước lượng như **Hình 23a**.



Hình 23a. Kết quả ước lượng mô hình đơn giản về lòng trung thành (với dữ liệu mới).

5.2 Giải thích các kết quả ước lượng mô hình PLS-SEM

Kết quả ước lượng mô hình PLS-SEM trên SmartPLS 4 gồm có năm loại kết quả khác nhau (như hộp bên trái **Hình 23**): Đồ thị (Graphical), Các kết quả cuối cùng (Final results), Các tiêu chí đánh giá chất lượng mô hình (Quality criteria), Thuật toán (Algorithm), và Mô hình và dữ liệu (Model and data). Thông thường ta chỉ sử dụng ba loại kết quả đầu tiên.

Kết quả xuất hiện **mặc định** trên sơ đồ (đồ thị) như ở **Hình 23, 23a** dựa hoàn toàn trên sơ đồ của mô hình đã được tạo sẵn trước đó (như ở **Hình 21**) với các giá trị hệ số ước lượng và giá trị R^2 được gán thêm vào. Các hệ số ước lượng trên các đường mũi tên ở **Hình 23, 23a** là các hệ số tải ngoài (outer loadings, **I**) và các hệ số đường dẫn (path coefficients, **p**). Các giá trị bên trong các hộp tròn là các hệ số R^2 của các phương trình hồi quy trong mô hình cấu trúc (cụ thể là 2 giá trị R^2 tương ứng với hai phương trình với các biến phụ thuộc/biến tiềm ẩn nội sinh là CUSA và CUSL). Hộp màu sẫm phía dưới bên trái là các lựa chọn được trực quan hóa trên sơ đồ. Ta có thể dễ dàng thay đổi các thống kê và các hệ số khác tùy vào mục đích nghiên cứu cụ thể.

Ví dụ, ta có thể thay đổi giá trị khác (như R^2 điều chỉnh, AVE, ρ_a , ρ_c , Cronbach's α) bằng cách nhấp vào “**Constructs**” trong “**Graphical output**”. Lưu ý, các thống kê như AVE, ρ_a , ρ_c , và Cronbach's α chỉ liên quan đến các biến tiềm ẩn/các cấu trúc gồm **nhiều biến quan sát** (cụ thể là các biến tiềm ẩn **COMP**, **LIKE**, và **CUSL** trong **Hình 23**). Riêng biến tiềm ẩn CUSA chỉ gồm một biến quan sát duy nhất nên hệ số luôn luôn bằng 1 và cũng không có các giá trị AVE, ρ_a , ρ_c , và Cronbach's α . Các thống kê này sẽ được trình bày ở phần sau. Bạn hãy tự thay đổi các thống kê khác nhau trong “**Constructs**”

trên kết quả ước lượng từ SmartPLS 4 của mình và xem các thay đổi trong các hộp hình tròn tương ứng với các biến tiềm ẩn/các cấu trúc.

Ngay phía dưới “**Constructs**” là “**Inner model**” (tức là mô hình cấu trúc) với các hệ số đường dẫn (**p**) được xuất hiện mặc định. Ta có thể thay đổi bằng các hệ số khác như hệ số tương quan (Correlations), hiệp phương sai (Covariances), ảnh hưởng gián tiếp (Indirect effects), tổng ảnh hưởng (Total effects), và f^2 . Thống kê f^2 là hệ số độ lớn ảnh hưởng (**Effect size**) là thước đo sự thay đổi của R^2 khi một biến ngoại sinh cụ thể bị loại bỏ sẽ có tầm quan trọng như thế nào đối với biến nội sinh trong mô hình cấu trúc. Nếu giá trị f^2 cao thì biến ngoại sinh đó có vai trò quan trọng trong mô hình cấu trúc, và ngược lại. Theo Hair và cộng sự (2022: trang 196), thì hệ số f^2 được tính như sau:

$$f^2 = \frac{R_{\text{included}}^2 - R_{\text{excluded}}^2}{1 - R_{\text{included}}^2} \quad (12)$$

Theo Cohen (1988), các giá trị f^2 lần lượt bằng 0.02, 0.15, và 0.35 cho biết các ảnh hưởng nhỏ, trung bình, và lớn của một biến tiềm ẩn ngoại sinh lên một biến tiềm ẩn nội sinh trong mô hình cấu trúc. Các giá trị f^2 nhỏ hơn 0.02 có nghĩa là biến tiềm ẩn ngoại sinh không có ảnh hưởng lên biến tiềm ẩn nội sinh.

Thường thì ta không thay đổi các lựa chọn đối với “**Outer model**” và “**Highlight paths**”. Cuối cùng, nếu ta chọn “**Blank**” cho tất cả các mô hình thì ta sẽ trở về mô hình chưa ước lượng ban đầu (như **Hình 21**).

Trong các kết quả cuối cùng (Final results) trình bày các kết quả ước lượng dưới dạng ma trận (**Matrix**) và bảng một chiều (**List**). Phần lớn những kết quả này đã được hoặc có thể được thể hiện một cách trực quan hóa trên đồ thị (như hệ số đường dẫn, hệ số tái ngoài, trọng số ngoài) nên chúng ít được sử dụng để trình bày kết quả ước lượng. Phổ biến nhất có lẽ là Đồ thị thanh (Bar chart) của các hệ số đường dẫn (như **Hình 24**), các giá trị ảnh hưởng gián tiếp (Indirect effects), và tổng ảnh hưởng (Total effects). Về các tiêu chí đánh giá chất lượng mô hình PLS-SEM sẽ được trình bày ở phần sau.

Cách giải thích các hệ số đường dẫn từ ước lượng **chuẩn hóa** sẽ như thế nào? Ví dụ, khi hồi quy Y_3 theo Y_1 và Y_2 trong mô hình cấu trúc, các hệ số hồi quy chuẩn hóa p_{13} và p_{23} giả sử có các giá trị lần lượt là 0.6 và 0.2 (và cả hai đều có ý nghĩa thống kê). Điều này chỉ ra rằng Y_1 có tầm quan trọng tương đối cao hơn trong việc giải thích Y_3 . Và hệ số p_{13} (=0.6) có nghĩa là: Nếu giữ nguyên biến tiềm ẩn Y_2 , và biến tiềm ẩn Y_1 được tăng thêm một đơn vị độ lệch chuẩn thì biến tiềm ẩn Y_3 tăng thêm 0.6 đơn vị độ lệch chuẩn. Hệ số p_{12} (=0.2) cũng được giải thích một cách tương tự. Áp dụng cho các kết quả ước lượng ở **Hình 23a** như sau. Các hệ số đường dẫn của mô hình cấu trúc cho thấy rằng **CUSA** có ảnh hưởng mạnh nhất lên **CUSL** ($p_{34} = 0.504$), kế tiếp là **LIKE** ($p_{24} = 0.342$) và **COMP** ($p_{14} = 0.009$). Hơn nữa, **LIKE** có ảnh hưởng mạnh hơn lên **CUSA** so với **COMP** bởi vì hệ số $p_{23} = 0.424$, trong khi hệ số $p_{13} = 0.162$. Theo nguyên tắc thông thường, đối với cỡ mẫu lên đến khoảng 500 quan sát, hệ số đường dẫn với các giá trị chuẩn hóa trên 0.20 thường có ý nghĩa thống kê và những hệ số đường dẫn có giá trị dưới 0.10 thường không có ý nghĩa thống kê (Hair và cộng sự, 2022: trang 101). Tuy nhiên, việc đưa ra các phát biểu chắc chắn về ý nghĩa thống kê của các hệ số đường dẫn

đòi hỏi phải xác định các sai số chuẩn của các hệ số ước lượng. Chúng ta sẽ bàn vấn đề này ở phần sau liên quan đến kỹ thuật mô phỏng (**Bootstrapping**).



Hình 24. Các hệ số đường dẫn của mô hình đơn giản về lòng trung thành⁵.

Các giá trị R^2 thường nằm trong khoảng giữa 0 và +1, và thể hiện lượng phương sai được giải thích trong một biến tiêm ẩn nội sinh. Ví dụ, giá trị R^2 bằng 0.7 cho biến tiêm ẩn nội sinh Y_3 (với hai biến tiêm ẩn ngoại sinh/hai biến giải thích Y_1 và Y_2) có nghĩa là 70% phương sai của Y_3 được giải thích bởi hai biến tiêm ẩn ngoại sinh Y_1 và Y_2 . Lưu ý rằng, với phương trình hồi quy có nhiều biến tiêm ẩn ngoại sinh/biến giải thích thì ta nên sử dụng R^2 điều chỉnh. Với kết quả trên **Hình 23a** thì $R^2 = 0.562$ có nghĩa là các biến tiêm ẩn **COMP**, **LIKE**, và **CUSA** cùng giải thích được 56.2% phương sai của biến tiêm ẩn nội sinh **CUSL**. Các biến tiêm ẩn **COMP** và **LIKE** cùng giải thích 29.5% phương sai của biến tiêm ẩn nội sinh **CUSA** ($R^2 = 0.295$).

6. ĐÁNH GIÁ MÔ HÌNH PLS-SEM

6.1 Giới thiệu

Theo Hair và cộng sự (2022: trang 110), việc đánh giá mô hình PLS-SEM thường theo quy trình hai bước như ở **Bảng 2**. Bước 1 là đánh giá các mô hình đo lường; và Bước 2 là đánh giá mô hình cấu trúc. Lưu ý rằng việc đánh giá mô hình đo lường kết quả (**Reflective measurement model**) và mô hình đánh giá nguyên nhân (**Formative measurement model**) là hoàn toàn khác nhau.

⁵ Ta có thể thay đổi đồ thị như đổi màu, cỡ chữ, xoay trực, đặt tên đồ thị, ... bằng cách chọn “Show properties”.

Bảng 2. Đánh giá các mô hình trong PLS-SEM.

A. Mô hình đo lường kết quả	B. Mô hình đo lường nguyên nhân
<ul style="list-style-type: none"> Độ tin cậy chỉ báo/Chất lượng biến quan sát (Indicator reliability) Độ tin cậy nhất quán nội bộ (Internal consistency reliability) [Hệ số α (Cronbach's alpha), độ tin cậy tổng hợp (Composite reliability) – ρ_c, hệ số tin cậy (Reliability coefficient) – ρ_a] Giá trị hội tụ (Convergent validity) [Phương sai trích trung bình – Average variance extracted] Giá trị phân biệt (Discriminant validity) 	<ul style="list-style-type: none"> Giá trị hội tụ (Convergent validity) Sự cộng tuyến giữa các biến quan sát (Collinearity between indicators) Mức ý nghĩa và sự phù hợp của các trọng số ngoài (Significance and relevance of outer weights)
C. Mô hình cấu trúc	
<ul style="list-style-type: none"> Sự cộng tuyến (Collinearity): VIF – Variance inflation factor (Hệ số khuyết đại phương sai) Mức ý nghĩa và sự phù hợp của các mối quan hệ mô hình cấu trúc (Significance and relevance of the structural model: Các hệ số đường dẫn – Path coefficients) Sức mạnh giải thích (Explanatory power): Hệ số xác định R^2 – Coefficient of determination Sức mạnh dự báo (Predictive power) So sánh mô hình (Model comparison) 	

Nguồn: Hair và cộng sự (2022: trang 100).

Theo Hair và cộng sự (2022: trang 111), logic của sử dụng nhiều biến chỉ báo (tức biến quan sát)⁶ so với chỉ một biến chỉ báo trong việc **đo lường cấu trúc** (tức biến tiềm ẩn) là **thước đo** (measure) sẽ chính xác hơn. Sự chính xác (accuracy) tăng lên bởi vì sử dụng nhiều biến chỉ báo để đo lường một khái niệm giúp giảm mức độ tổng thể của sai số đo lường (**measurement error**) vốn gắn liền với các biến chỉ báo. Hơn nữa, **sự đo lường**⁷ bằng nhiều biến chỉ báo/nhiều thang đo (multi-item measurement) có khả năng cao hơn trong việc thể hiện nhiều khía cạnh khác nhau của một khái niệm, nhất là các khái niệm phức tạp như lòng tin (trust), sự hài lòng (satisfaction), sự cam kết (commitment), cảm nhận (perception), v.v. Tuy nhiên, thậm chí khi sử dụng nhiều biến chỉ báo khác nhau để đo lường các khái niệm thì sự đo lường vẫn rất có thể tồn tại một

⁶ Nhắc lại: Có nhiều tên gọi trong tiếng Anh như Observed variable/Observed manifest/Indicator/Item.

⁷ Theo Malhotra (2020: trang 268) và Nunan và cộng sự (2020: trang 329) thì “Đo lường có nghĩa là gán các con số hoặc các ký hiệu khác cho các đặc điểm của các đối tượng theo các quy tắc nhất định được chỉ định trước. Chúng ta không đo lường đối tượng, mà là một đặc điểm nào đó của đối tượng” Các đối tượng (objects) là những người tiêu dùng, nhà sản xuất; còn các đặc điểm (characteristics) của đối tượng là cảm nhận (perceptions), thái độ (attitudes), sở thích (preferences), hay các đặc điểm liên quan khác. Nunan và cộng sự (2020: trang 354) cho rằng một sự đo lường không phải là một giá trị thực sự của đặc điểm quan tâm mà là một quan sát của nó.

mức độ sai số đo lường nhất định. Có nhiều nguồn gây ra sai số đo lường trong các nghiên cứu khoa học xã hội như sử dụng từ ngữ kém cỏi cho các câu hỏi khảo sát, hiểu sai phương pháp sử dụng thang đo, và áp dụng một phương pháp thống kê không đúng đắn. Tất cả các lý do này có thể dẫn tới các sai số ngẫu nhiên (**random error**) và sai số hệ thống (**systematic error**).

Theo Molhotra (2020: trang 302) và Nunan và cộng sự (2020: trang 352), mô hình điểm số đúng (**true score model**) cung cấp một khung phân tích để hiểu sự chính xác trong đo lường. Theo mô hình này thì:

$$X_O = X_T + E_S + E_R \quad (13)$$

Trong đó:

X_O = Điểm số/giá trị quan sát hay giá trị được đo lường

X_T = Điểm số/giá trị đúng của một thuộc tính

E_S = Sai số hệ thống

E_R = Sai số ngẫu nhiên

Tổng sai số đo lường bao gồm sai số hệ thống, E_S , và sai số ngẫu nhiên, E_R . **Sai số hệ thống**⁸ ảnh hưởng sự đo lường theo một cách cố định. Nó thể hiện các yếu tố ổn định ảnh hưởng điểm số quan sát theo một cách giống nhau mỗi lần sự đo lường được thực hiện chẳng hạn như các yếu tố cơ học (mechanical factors). Như Mooi và cộng sự (2018: trang 38) đưa ra ví dụ là nếu ta yêu cầu những khách hàng đánh giá về quảng cáo trên TV và trả thù lao cho họ, thì họ có thể cung cấp thông tin tốt hơn họ nghĩ. Điều này làm cho ta nghĩ rằng quảng cáo trên TV là thứ vị một cách có hệ thống hơn vốn dĩ trên thực tế. Ngược lại, **sai số ngẫu nhiên** thì không cố định. Nó thể hiện các yếu tố nhất thời (transient factors) ảnh hưởng điểm số quan sát theo các cách khác nhau mỗi lần sự đo lường được thực hiện như các yếu tố thuộc về cá nhân nhất thời (transient personal factors) như cảm xúc của người trả lời câu hỏi khảo sát, sự mệt mỏi, ... và các yếu tố ngữ cảnh (situational factors) như mức độ tiếng ồn lúc trả lời câu hỏi khảo sát, có hay không có sự hiện diện của người khác trong lúc phỏng vấn, ... Ví dụ, một số khách hàng có thể có một ngày tốt lành và chỉ ra rằng họ thích quảng cáo trên TV; trong khi đó, một số lại có một ngày tồi tệ thì lại trả lời ngược lại (Mooi và cộng sự, 2018: trang 38). Một vài yếu tố ảnh hưởng sai số đo lường được liệt kê trong **Bảng 3**.

Phân biệt giữa sai số hệ thống và sai số ngẫu nhiên là vô cùng quan trọng để hiểu khái niệm độ tin cậy/độ ổn định (**reliability**) và độ chuẩn xác/độ xác thực (**validity**). Để hiểu rõ hơn về các thành phần của hai khái niệm này, hãy tham khảo các tài liệu như Malhotra (2020: Chapters 8 & 9); Nunan và cộng sự (2020: Chapter 12). Hair và cộng sự (2022: trang 111) cho rằng **sai số hệ thống ảnh hưởng đến độ chuẩn xác, còn sai số ngẫu nhiên**

⁸ Mooi và cộng sự (2018: trang 38) cho rằng **sai số hệ thống** làm cho đo lường thực tế cao hơn hay thấp hơn một cách nhất quán so với giá trị thật của nó; còn **sai số ngẫu nhiên** gây ra sự biến thiên (ngẫu nhiên) giữa thứ ta thật sự đo và thứ ta muốn đo.

ảnh hưởng đến độ tin cậy. Theo Mooi và cộng sự (2018: trang 38), thì **độ chuẩn xác** đề cập đến việc liệu chúng ta có đang đo lường những gì mà chúng ta muốn đo lường hay không và đó là trường hợp mà sai số hệ thống ***Es*** là không đáng kể. **Độ tin cậy** là mức độ mà những gì chúng ta đo lường không có sai số ngẫu nhiên và do đó liên quan đến trường hợp ***Er*** bằng không.

Bảng 3. Các yếu tố ảnh hưởng sai số đo lường.

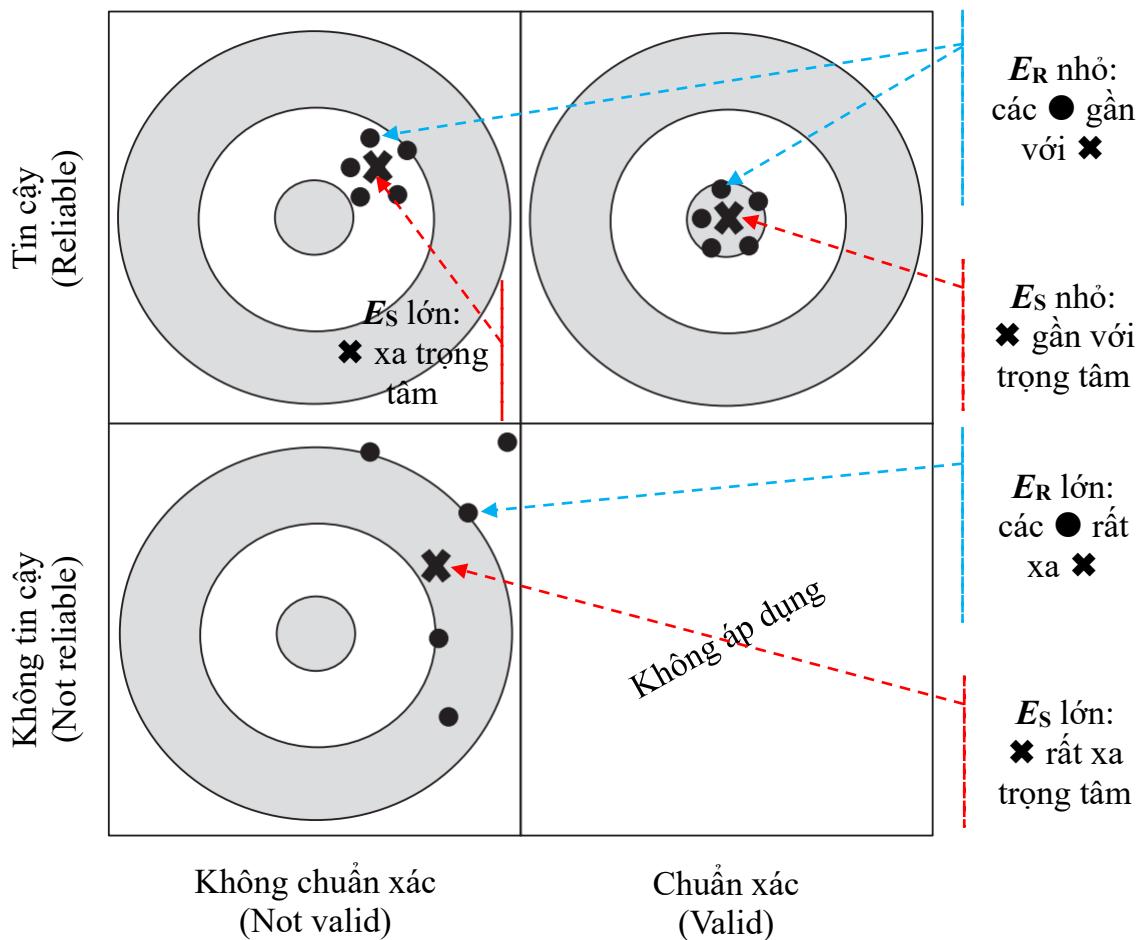
1. Các đặc điểm tương đối ổn định của cá nhân ảnh hưởng đến điểm số bài kiểm tra/giá trị quan sát, chẳng hạn như sự thông minh, sự mong muốn xã hội (xu hướng đưa câu trả lời làm cho mình trông tốt hơn/xấu đi với phỏng vấn viên) và trình độ giáo dục.
2. Các yếu tố cá nhân ngắn hạn hoặc nhất thời, chẳng hạn như sức khỏe, cảm xúc, sự mệt mỏi.
3. Các yếu tố tình huống/bối cảnh, chẳng hạn như sự hiện diện của người khác, mức độ tiếng ồn và sự phiền nhiễu.
4. Lấy mẫu các hạng mục được đưa vào thang đo (nghĩa là số biến quan sát cho một khái niệm/cấu trúc): thêm, xóa hoặc thay đổi các hạng mục thang đo.
5. Thiếu sự rõ ràng của thang đo, bao gồm các hướng dẫn hoặc chính các hạng mục.
6. Các yếu tố cơ học, chẳng hạn như chất lượng in ấn kém, quá nhiều các hạng mục trong bảng câu hỏi, và thiết kế bảng câu hỏi kém.
7. Quản lý thang đo, chẳng hạn như sự khác biệt giữa những người phỏng vấn.
8. Các yếu tố phân tích, chẳng hạn như sự khác biệt trong cách tính điểm và phân tích thống kê.

Nguồn: Malhotra (2020: trang 302), Nunan và cộng sự (2020: trang 353).

Hình 25 giải thích sự khác biệt giữa độ tin cậy và độ chuẩn xác. Năm dấu chấm tròn (●) thể hiện cho năm lần đo lường (ví dụ như 5 biến chỉ báo về sự hài lòng của khách hàng về một sản phẩm hoặc dịch vụ cụ thể), và giá trị trung bình của năm lần đo lường được thể hiện bằng dấu chéo (✖). Vòng tròn trung tâm thể hiện giá trị thật hay điểm số đúng về sự hài lòng của khách hàng. **Độ chuẩn xác** (validity) được thể hiện khi dấu chéo (✖) gần với vòng tròn trung tâm. Giá trị trung bình của năm lần đo càng gần với giá trị thật, thì sai số hệ thống càng thấp và độ chuẩn xác càng cao. **Độ tin cậy** (reliability) là các khoảng cách giữa các dấu tròn (●). Nếu sai số ngẫu nhiên thấp, tất cả các dấu tròn gần với nhau, thì thước đo càng tin cậy, cho dù các dấu tròn không nhất thiết phải gần vòng tròn trung tâm. Theo Hair và cộng sự (2022: trang 112) và Mooi và cộng sự (2018: trang 39) độ tin cậy là điều kiện cần cho độ chuẩn xác. Và chính vì thế mà trường hợp “không tin cậy nhưng chuẩn xác” là không thể xảy ra.

Theo Hair và cộng sự (2022: trang 112-13), khi đánh giá các mô hình đo lường, chúng ta phải phân biệt giữa các cấu trúc được đo lường theo kết quả và theo nguyên nhân. Hai cách tiếp cận dựa trên các khái niệm khác nhau và do đó đòi hỏi phải xem xét các biện pháp đánh giá khác nhau. **Mô hình đo lường kết quả** được đánh giá dựa trên độ tin cậy của các chỉ báo (indicator reliability), độ tin cậy nhất quán bên trong (internal consistency reliability), giá trị hội tụ (convergent validity) và giá trị phân biệt. Các tiêu chí đánh giá mô hình đo lường kết quả không thể được áp dụng hoàn toàn cho **mô hình**

đo lường nguyên nhân. Với các thước đo nguyên nhân, bước đầu tiên là đảm bảo giá trị nội dung (content validity) trước khi thu thập dữ liệu và ước lượng mô hình đường dẫn PLS. Sau khi ước lượng mô hình, các tiêu chí khác nhau được sử dụng để đánh giá các thước đo nguyên nhân về giá trị hội tụ, sự cộng tuyến giữa các biến chỉ báo⁹, và sự ý nghĩa thống kê cũng như sự phù hợp của các trọng số của biến chỉ báo (tức là các trọng số ngoài, *w*). Lưu ý rằng, các tiêu chí đánh giá các mô hình đo lường kết quả và nguyên nhân không thể áp dụng cho các cấu trúc hình thành từ chỉ một biến chỉ báo duy nhất. Trong một cấu trúc đơn biến chỉ báo thì mối quan hệ giữa cấu trúc và biến quan sát luôn là 1 vì điểm số/giá trị đo lường của cấu trúc cũng chính là điểm số/giá trị đo lường của biến quan sát hình thành nên cấu trúc.



Hình 25. Phân biệt độ chuẩn xác và độ tin cậy (Hair và cộng sự, 2022: trang 112).

Các giá trị ước lượng của **mô hình cấu trúc** không được phân tích cho đến khi độ tin cậy và độ chuẩn xác của các cấu trúc đã được thiết lập. Nếu sự đánh giá của các mô hình đo lường kết quả và/hoặc mô hình đo lường nguyên nhân cung cấp bằng chứng về chất lượng của các thước đo (cấu trúc), thì các ước lượng của mô hình cấu trúc mới được đánh giá. Đánh giá mô hình cấu trúc thường liên quan đến sự cộng tuyến tiềm năng giữa các cấu trúc, kiểm định độ chuẩn xác danh nghĩa (**nomological validity**) của

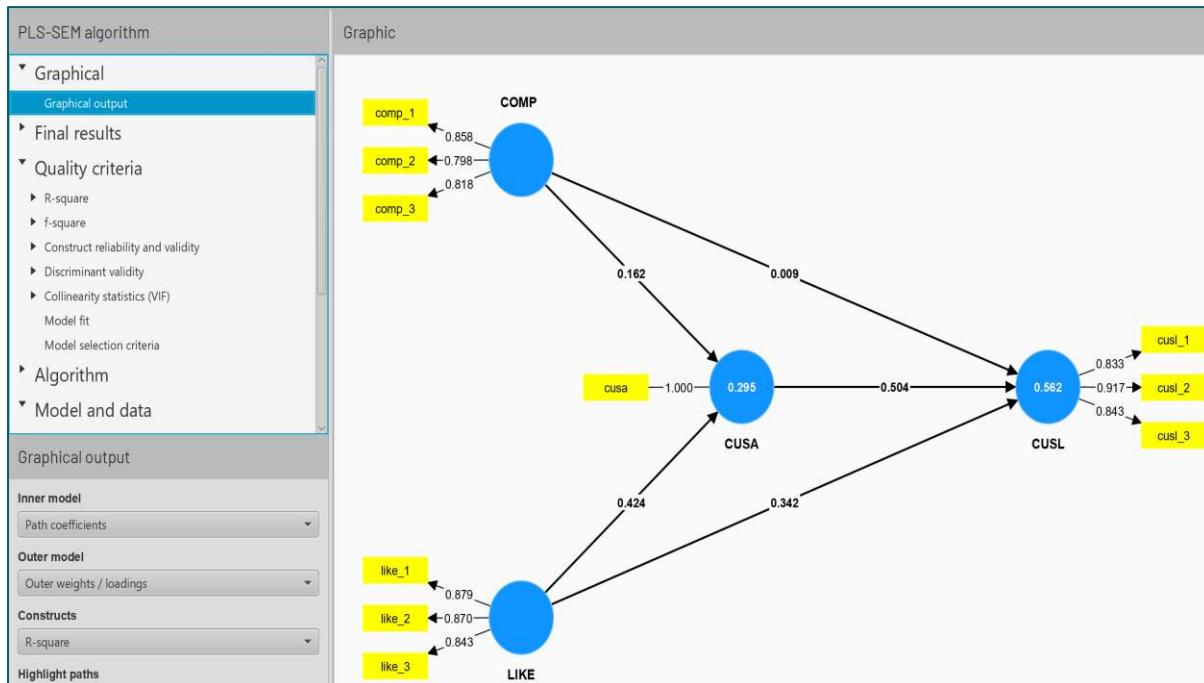
⁹ Vì thước đo nguyên nhân hay cấu trúc nguyên nhân là một hàm tuyến tính của các biến chỉ số như phương trình (2) nên vấn đề đa cộng tuyến cần phải đặc biệt quan tâm.

mô hình và khả năng mà mô hình giải thích và dự đoán phương sai của các biến phụ thuộc tiềm ẩn (các biến tiềm ẩn nội sinh). Kiểm định độ chuẩn xác danh nghĩa liên quan đến đánh giá độ lớn và mức ý nghĩa của các hệ số đường dẫn, trong khi sức mạnh giải thích và dự đoán được đánh giá dựa vào hệ số xác định (R^2), độ lớn ảnh hưởng (f^2), và các kết quả của thuật toán PLS_{predict}.

6.2 Đánh giá mô hình đo lường kết quả

Theo Hair và cộng sự (2022: trang 117-118), đánh giá các mô hình đo lường kết quả bao gồm đánh giá độ tin cậy của các thước đo, cả ở mức biến chỉ báo (độ tin cậy chỉ báo/indicator reliability) và cấu trúc (độ tin cậy nhất quán nội bộ/internal consistency reliability). Đánh giá độ chuẩn xác tập trung vào việc đánh giá hai loại giá trị (validity types). Thứ nhất là giá trị hội tụ (convergent validity) của mỗi thước đo bằng cách sử dụng phương sai trích trung bình (AVE). Thứ hai là giá trị phân biệt (discriminant validity) trong đó ta so sánh tất cả các thước đo cấu trúc trong cùng mô hình dựa trên chỉ số tương quan HTMT (Heterotrait-monotrait ratio of correlations). Các nội dung đánh giá mô hình đo lường kết quả được minh họa ở **Hình 26**.

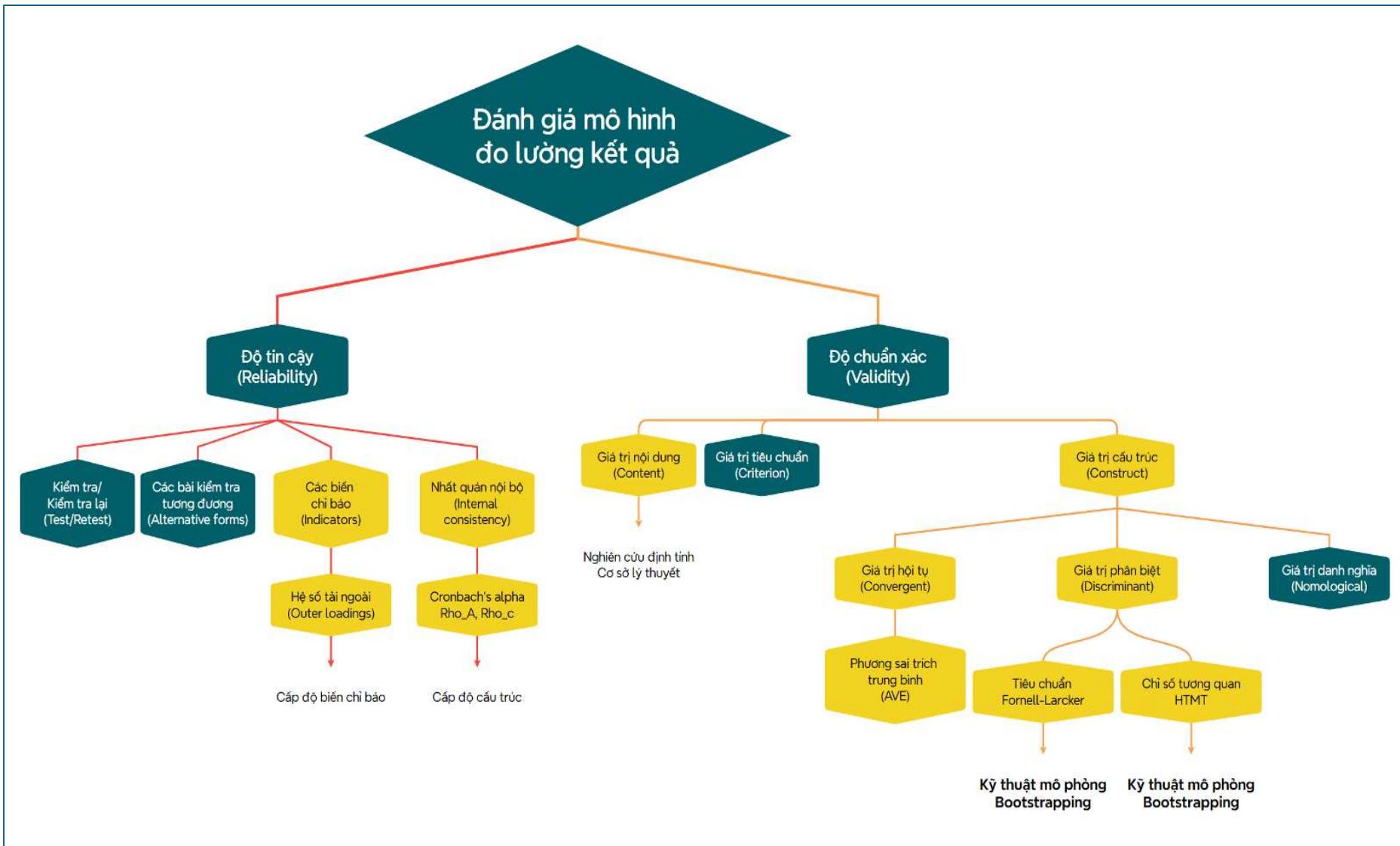
Để minh họa cho việc đánh giá mô hình đo lường, ta tiếp tục sử dụng kết quả ước lượng từ mô hình đơn giản như ở **Hình 23b**.



Hình 23b. Kết quả ước lượng mô hình đơn giản về lòng trung thành (với dữ liệu mới).

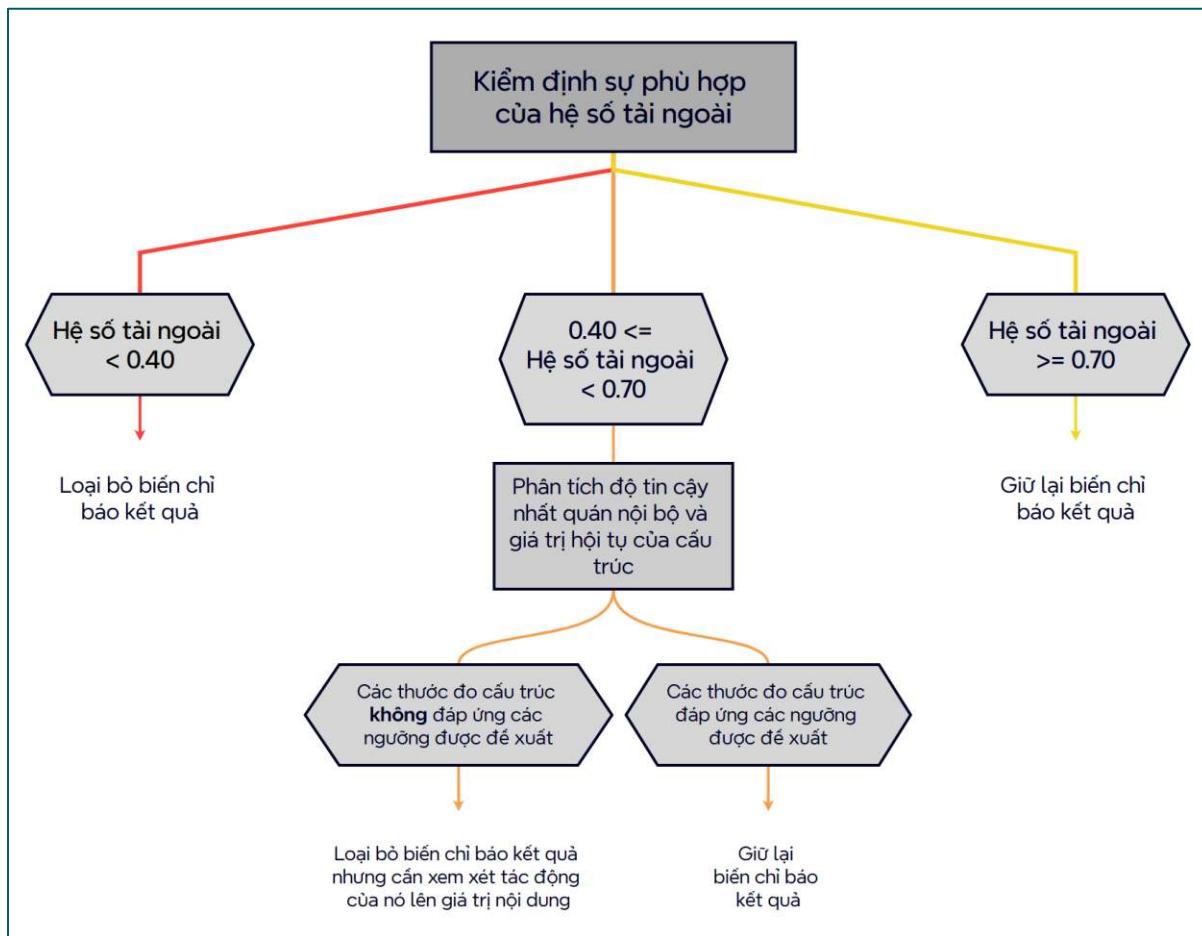
Bước 1: Đánh giá độ tin cậy của biến chỉ báo (Indicator reliability).

Bước đầu tiên trong đánh giá mô hình đo lường kết quả liên quan đến việc xem xét các hệ số tải ngoài của các biến chỉ báo (*I*) như ở các phương trình (3) đến (8). Các hệ số tải ngoài theo một cấu trúc cao cho biết các biến chỉ báo liên quan có nhiều điểm chung.



Hình 26. Nội dung đánh giá mô hình đo lường kết quả.

Theo Hair và cộng sự (2022: trang 117), độ lớn của hệ số tải ngoài được gọi chung là độ tin cậy chỉ báo. Ở mức tối thiểu, các hệ số tải ngoài của tất cả các biến chỉ báo phải có ý nghĩa thống kê. Tuy nhiên, một hệ số tải ngoài có ý nghĩa thống kê có lẽ vẫn còn khá yếu, nên một quy tắc chung là các hệ số tải ngoài chuẩn hóa (standardized outer loadings) phải từ **0.708** hoặc cao hơn. Lý do đưa ra con này là bởi vì bình phương của hệ số tải ngoài chuẩn hóa thể hiện mức độ tương quan của một biến chỉ báo với tất cả các biến chỉ báo khác (gọi là **communality**); nghĩa là, nó thể hiện bao nhiêu phần trăm biến thiên trong một biến chỉ báo được giải thích bởi cấu trúc và được gọi là phương sai trích từ một biến chỉ báo. Và quy tắc chung này có nghĩa là một biến tiềm ẩn (tức cấu trúc) nên giải thích một phần đáng kể cho phương sai của mỗi biến chỉ báo, tối thiểu thường là **50%** ($0.708^2 = 0.5$). Việc giải thích ít nhất 50% phương sai của một biến chỉ báo hàm ý rằng phương sai được chia sẻ giữa cấu trúc và biến chỉ báo của nó phải cao hơn sai số đo lường. Trên thực tế thì hệ số tải ngoài khoảng **0.7** là có thể chấp nhận được.



Hình 27. Quy trình kiểm tra độ tin cậy chỉ báo (Hair và cộng sự, 2022: trang 118).

Về nguyên tắc, biến chỉ báo nào có hệ số tải **nhỏ hơn 0.7** thì nên bị loại bỏ ra. Tuy nhiên, người nghiên cứu nên xem xét thận trọng ảnh hưởng của việc loại bỏ biến chỉ báo đó lên các thước đo độ tin cậy và độ chuẩn xác khác. Nói chung, các biến chỉ báo có hệ số tải ngoài **giữa 0.4 và 0.7** thì nên được xem xét loại bỏ chỉ khi việc loại bỏ biến chỉ báo dẫn đến một sự **giả tăng** trong **độ tin cậy nhất quán nội bộ** hoặc **giá trị hội tụ**.

Các biến chỉ báo với hệ số tải ngoài yếu đuối khi được giữ lại trên cơ sở đóng góp của chúng cho giá trị nội dung (**content validity**). Các biến chỉ báo với hệ số tải ngoài rất thấp (**0.4 <**) luôn luôn bị loại bỏ ra khỏi cấu trúc (Hair và cộng sự, 2011). Quy trình kiểm tra độ tin cậy chỉ báo được minh họa ở **Hình 27**.

Trở lại ví dụ của chúng ta thì hầu hết các hệ số tải ngoài (các con số ướng với các biến chỉ báo trong hộp chữ nhật màu vàng trên **Hình 23b**) đều lớn hơn 0.708. Điều này gợi ý rằng các biến chỉ báo đều thỏa mãn điều kiện về độ tin cậy chỉ báo. Biến chỉ báo có độ tin cậy chỉ báo thấp nhất là biến **comp_2** với giá trị là 0.637 (= 0.798²), và biến chỉ báo có độ tin cậy chỉ báo cao nhất là biến **cusl_2** với giá trị là 0.841 (= 0.917²). Một cách khác để liệt kê các hệ số tải ngoài như sau: **Final results → Outer loadings** (như minh họa trên **Hình 28**).

	COMP	CUSA	CUSL	LICK
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

Hình 27. Hệ số tải ngoài của các biến chỉ báo từ mô hình đơn giản.

Bước 2: Đánh giá độ tin cậy nhất quán nội bộ (Internal consistency reliability).

Theo Hair và cộng sự (2022: trang 118), tiêu chí truyền thống để đánh giá độ tin cậy nhất quán nội bộ là sử dụng hệ số **Cronbach's α** . Hệ số này được định nghĩa như sau:

$$\text{Cronbach's } \alpha = \left(\frac{M}{M-1} \right) \cdot \left(1 - \frac{\sum_{i=1}^M s_i^2}{s_t^2} \right) \quad (14)$$

Trong công thức này, s_i^2 là phương sai của biến chỉ báo i của một cấu trúc cụ thể, được đo lường với M biến chỉ báo ($i = 1, \dots, M$), và s_t^2 là phương sai của tổng tất cả M biến chỉ báo của cấu trúc đó.

Hệ số **Cronbach's α** dao động từ 0 và 1, với các giá trị càng cao thì mức độ tin cậy của cấu trúc càng cao. Cụ thể, các giá trị **từ 0.60 đến 0.70** là có thể chấp nhận được trong

nghiên cứu có tính khám phá (**exploratory research**), trong khi trong các giai đoạn cao hơn của nghiên cứu, thì các giá trị **giữa 0.70 và 0.90** có thể được coi là thỏa đáng. Các giá trị **trên 0.90** (và tuyệt đối trên 0.95) thì không mong muốn bởi vì chúng điển hình là kết quả của sự dư thừa các biến quan sát (**redundant items**); nghĩa là gần như lặp lại cùng một câu hỏi khi phỏng vấn. Sau cùng, các giá trị dưới 0.60 thể hiện không đạt độ tin cậy nhất quán nội bộ. Tuy nhiên, một điểm yếu của hệ số **Cronbach's α** là nó giả định rằng tất cả các biến chỉ báo đều tin cậy như nhau (tức là, tất cả các biến chỉ báo có các hệ số tải ngoài bằng nhau trong cấu trúc). Hơn nữa, hệ số **Cronbach's α** nhạy cảm với số biến chỉ báo trong một cấu trúc và có xu hướng ước lượng thấp độ tin cậy nhất quán nội bộ. Vì thế, hệ số **Cronbach's α** có thể được sử dụng như một thước đo bảo thủ hơn của độ tin cậy nhất quán nội bộ ([Hair và cộng sự \(2022: trang 119\)](#).

Do hạn chế của hệ số **Cronbach's α** nên cần phải áp dụng một thước đo khác cho độ tin cậy nhất quán nội bộ; đó là, độ tin cậy tổng hợp, ký hiệu là **ρ_c (composite reliability)**. Hệ số này có xem xét đến các hệ số tải ngoài khác nhau của các biến chỉ báo và được tính theo công thức sau:

$$\rho_c = \frac{\left(\sum_{i=1}^M l_i\right)^2}{\left(\sum_{i=1}^M l_i\right)^2 + \sum_{i=1}^M var(e_i)} \quad (15)$$

Trong công thức này, l_i là hệ số tải ngoài chuẩn hóa của biến chỉ báo i của một cấu trúc cụ thể, được đo lường với M biến chỉ báo ($i = 1, \dots, M$), e_i là sai số đo lường của biến chỉ báo i , và $var(e_i)$ là phương sai của sai số đo lường được định nghĩa là $1 - l_i^2$.

Tương tự hệ số **Cronbach's α** , hệ số **ρ_c** dao động từ 0 và 1, với các giá trị càng cao thì mức độ tin cậy của cấu trúc càng cao. Cụ thể, các giá trị **từ 0.60 đến 0.70** là có thể chấp nhận được trong nghiên cứu có tính khám phá (**exploratory research**), trong khi trong các giai đoạn cao hơn của nghiên cứu, thì các giá trị **giữa 0.70 và 0.90** có thể được coi là thỏa đáng. Các giá trị **trên 0.90** (và tuyệt đối trên 0.95) thì không mong muốn bởi vì chúng điển hình là kết quả của sự dư thừa các biến quan sát (**redundant items**); nghĩa là gần như lặp lại cùng một câu hỏi khi phỏng vấn. Sau cùng, các giá trị dưới 0.60 thể hiện không đạt độ tin cậy nhất quán nội bộ ([Hair và cộng sự \(2022: trang 119\)](#).

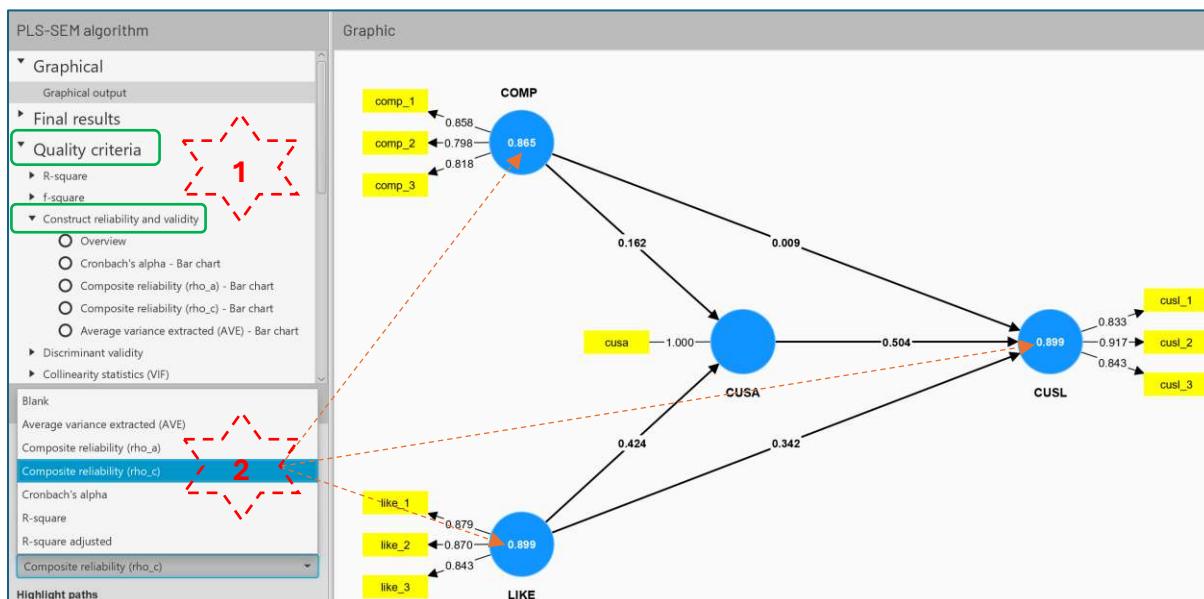
Trong khi hệ số **Cronbach's α** có xu hướng bảo thủ thì hệ số **ρ_c** có lẽ hơi quá tự do. Cho nên một hệ số khác về độ tin cậy được giới thiệu; đó là hệ số **ρ_A (exact or consistent reliability coefficient)**. Hệ số này có công thức như sau:

$$\rho_A = (\hat{w}' \hat{w})^2 \cdot \frac{\hat{w}'(S - \text{diag}(S))\hat{w}}{\hat{w}'(\hat{w} \hat{w}' - \text{diag}(\hat{w} \hat{w}'))\hat{w}} \quad (16)$$

Trong công thức này, \hat{w}_i là các giá trị ước lượng của trọng số ngoài, **diag** là đường chéo của ma trận tương ứng, và **S** là ma trận hiệp phương sai mẫu.

Hệ số ρ_A thường có giá trị nằm giữa hệ số **Cronbach's α** và hệ số tin cậy tổng hợp ρ_c , và vì thế được xem là một sự thỏa hiệp tốt giữa hai thước đo này (Hair và cộng sự, 2019; 2022: trang 119). Chính vì thế, cách quyết định về độ tin cậy nhất quán nội bộ với hệ số tin cậy ρ_A cũng tương tự như với hệ số **Cronbach's α** và hệ số tin cậy tổng hợp ρ_c .

Trở lại ví dụ về mô hình đơn giản của chúng ta. Trên SmartPLS 4 có hai cách để thể hiện các hệ số về độ tin cậy nhất quán nội bộ như ở **Hình 28**. Thứ nhất, trong hộp bên trái phía trên, chọn **Quality criteria → Construct reliability and validity → Overview**, để xem các hệ số dưới dạng bảng. Hoặc có thể chọn xem từng hệ số dưới dạng đồ thị Bạn hãy tự thực hiện với kết quả ước lượng của mình. Thứ hai, trong “**Constructs**” ở hộp bên trái phía dưới (tức **Graphical output**), thay “R-square” (mặc định) thành một trong ba loại hệ số khác (như **Hình 28**). Với cách này thì giá trị các hệ số sẽ hiện diện trong hộp tròn của các cấu trúc (biến tiềm ẩn) có nhiều hơn một biến chỉ báo. Từ kết quả này ta nhận thấy tất cả (3) cấu trúc đều có mức cao và thỏa đáng về độ tin cậy nhất quán nội bộ (trong khoảng 0.7 đến 0.9). Tuy nhiên, cấu trúc **CUSA** không xuất hiện giá trị ρ_c (hoặc ρ_A hoặc **Cronbach's α**) bởi vì nó chỉ gồm một biến chỉ báo duy nhất.



Hình 28. Độ tin cậy nhất quán nội bộ của các cấu trúc đa chỉ báo.

Bước 3: Dánh giá giá trị hội tụ (Convergent validity).

Theo Hair và cộng sự (2022: trang 118), giá trị hội tụ là mức độ mà một thước đo (**measure**) tương quan **dương** với các thước đo khác của cùng cấu trúc. **Thước đo ở đây chính là biến chỉ báo**. Các biến chỉ báo của một cấu trúc kết quả (**reflective construct**) cụ thể phải hội tụ hoặc chia sẻ một tỷ lệ cao giá trị phuơng sai. Nhớ lại rằng, bình phuơng của một hệ số tải ngoài của một biến chỉ báo i (L_i^2) thể hiện phuơng sai chia sẻ (chung) của biến chỉ báo đó với một cấu trúc kết quả cụ thể, mà cấu trúc này đại diện cho tất cả các biến chỉ báo liên quan, trong đó có biến chỉ báo i .

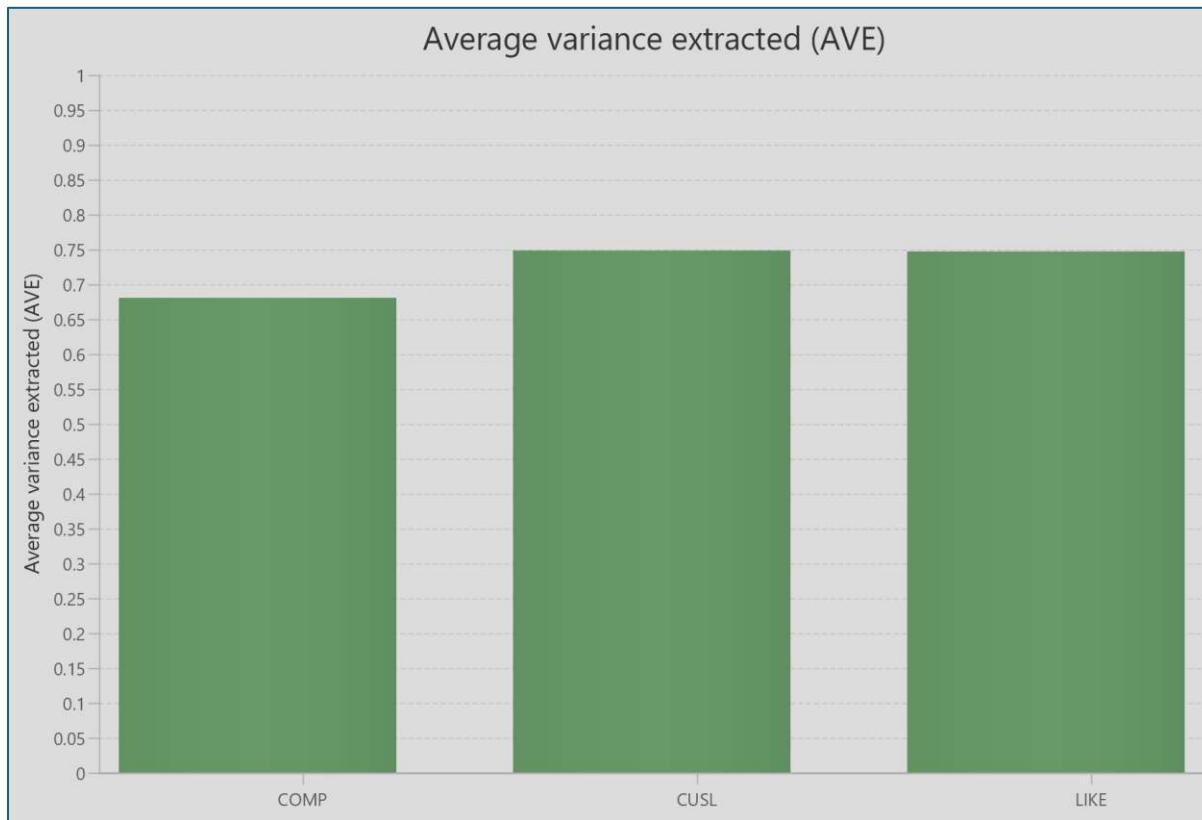
Một thước đo phổ biến để thiết lập giá trị hội tụ ở mức độ cấu trúc là phương sai trích trung bình (**AVE** – average variance extracted). **AVE** là tương đương với phương sai chung (**communality**) của một cấu trúc. Tiêu chí này được tính như sau:

$$AVE = \left(\frac{\sum_{i=1}^M l_i^2}{M} \right) \quad (17)$$

Trong đó, l_i là hệ số tải ngoài chuẩn hóa của biến chỉ báo i của một cấu trúc kết quả cụ thể với M biến chỉ báo.

Cách áp dụng giống với logic của độ tin cậy chỉ báo. Nghĩa là, một giá trị **AVE bằng 0.50 hoặc cao hơn** chỉ ra rằng, trung bình, cấu trúc giải thích hơn một nửa phương sai của các biến chỉ báo của nó. Ngược lại, một giá trị **AVE nhỏ hơn 0.50** chỉ ra rằng, trung bình, nhiều phương sai còn chứa trong sai số của các biến chỉ báo hơn trong phương sai được giải thích bởi cấu trúc.

Trở lại ví dụ về mô hình đơn giản của chúng ta. Để có kết quả về các giá trị **AVE**, ta có thể thực hiện một trong hai cách như ở **Hình 28**. Theo cách thứ nhất với lựa chọn là biểu diễn trên đồ thị, ta có kết quả như ở **Hình 29**. Các giá trị **AVE** của **COMP** (0.681), **CUSL** (0.748), và **LIKE** (0.747) đều cao hơn mức yêu cầu tối thiểu 0.50. Như vậy, các biến chỉ báo của cả ba cấu trúc đều có các mức giá trị hội tụ cao.

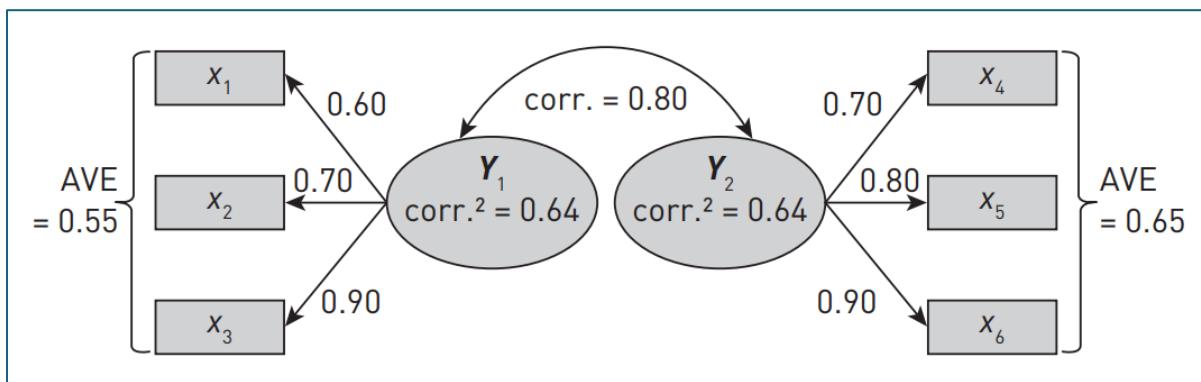


Hình 29. Phương sai trích trung bình của các cấu trúc trong mô hình đơn giản.

Bước 4: Đánh giá giá trị phân biệt (Discriminant validity).

Theo Hair và cộng sự (2022: trang 120), giá trị phân biệt là mức độ mà một cấu trúc thực sự khác biệt với các cấu trúc khác theo các tiêu chuẩn thực nghiệm. Vì thế, thiết lập giá trị phân biệt hàm ý rằng một cấu trúc là duy nhất và chứa các thông tin/khía cạnh không được thể hiện bởi các cấu trúc khác trong mô hình. Thường thì các nhà nghiên cứu sử dụng tiêu chí Fornell-Larcker để đánh giá giá trị phân biệt. Tiêu chí này so sánh căn bậc hai của các giá trị AVE với các hệ số tương quan của biến tiềm ẩn. Cụ thể, căn bậc hai của AVE của mỗi cấu trúc phải lớn hơn hệ số tương quan cao nhất của nó với bất kỳ cấu trúc nào khác.

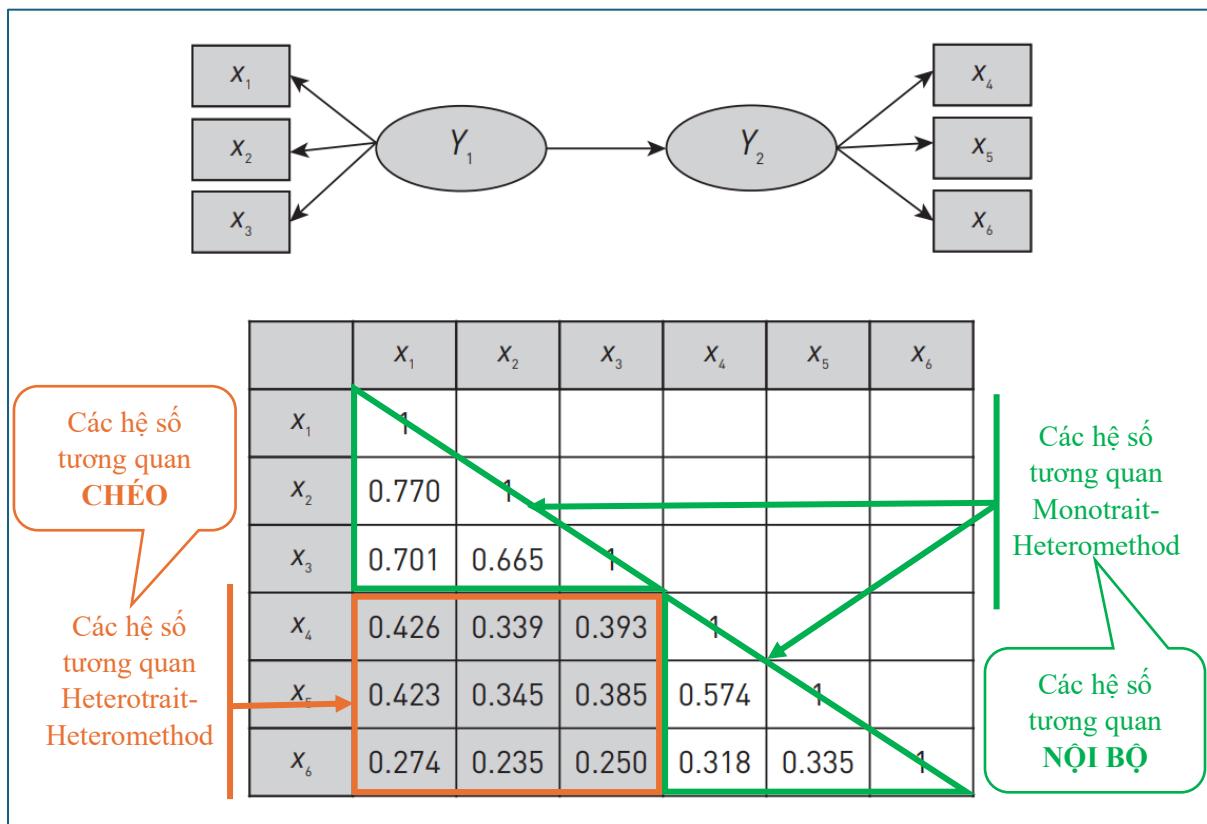
Xét một ví dụ như ở **Hình 30** (Hair và cộng sự, 2022: trang 120). Các giá trị AVE của các cấu trúc Y_1 và Y_2 lần lượt là 0.55 và 0.65. Các giá trị AVE này thu được bằng cách bình phương từng hệ số tải ngoài, tính tổng của ba hệ số tải ngoài bình phương, và sau đó tính toán giá trị trung bình. Ví dụ: đối với cấu trúc Y_1 thì bình phương của các hệ số tải ngoài 0.60, 0.70 và 0.90 lần lượt là 0.36, 0.49 và 0.81. Tổng của ba con số này là 1.66 và giá trị trung bình là 0.55 (tức là 1.66/3). Hệ số tương quan giữa các cấu trúc Y_1 và Y_2 (như được chỉ ra bởi mũi tên hai đầu liên kết hai cấu trúc) là 0.80. Bình phương của hệ số tương quan 0.80 (tức $0.80^2 = 0.64$) chỉ ra rằng 64% của biến thiên của mỗi cấu trúc được giải thích bởi một cấu trúc khác. Do đó, Y_1 giải thích ít phương sai hơn trong các biến chỉ báo X_1 , X_2 và X_3 so với Y_2 ; hàm ý rằng hai cấu trúc Y_1 và Y_2 — *mặc dù khác nhau về mặt khái niệm* — không đủ khác nhau theo các tiêu chí Fornell-Larcker. Do đó, giá trị phân biệt trong ví dụ này không được thiết lập.



Hình 30. Minh họa giá trị phân biệt (Hair và cộng sự, 2022: trang 121).

Tuy nhiên, tiêu chí Fornell-Larcker bị các nhà nghiên cứu nghi ngờ về tính hiệu lực. Henseler và cộng sự (2015) cho thấy tiêu chí này thể hiện rất kém khi các hệ số tải ngoài của các biến chỉ báo trong các cấu trúc đang được xem xét chỉ khác nhau một chút (ví dụ: tất cả các hệ số tải ngoài dao động trong khoảng từ 0.60 đến 0.80) — như thường xảy ra trong các ứng dụng thực nghiệm của PLS-SEM. Khi các hệ số tải ngoài của biến chỉ báo thay đổi một cách mạnh mẽ hơn, kết quả thực hiện theo tiêu chí Fornell-Larcker trong việc phát hiện các vấn đề giá trị phân biệt được cải thiện, nhưng xét về tổng thể vẫn rất kém. Vì vậy, tiêu chí Fornell-Larcker không thực sự đánh giá được tính phân biệt của một thang đo.

Henseler và cộng sự (2015) đề xuất phương pháp đánh giá thay thế và được chấp nhận rộng rãi trong giới nghiên cứu, gọi là chỉ số tương quan Heterotrait-Monotrait, viết tắt là **HTMT**. Chỉ số HTMT được tính toán dựa trên nền tảng ma trận đa đặc điểm – đa phương pháp (multitrait-multimethod matrix – **MTMM**) do Campbell và Fiske (1959) đề xuất. Ma trận này yêu cầu có ít nhất hai tập các biến chỉ báo để đo lường hai cấu trúc/hai biến tiềm ẩn khác nhau và mỗi tập các biến chỉ báo có số lượng bằng nhau (ma trận cân đối). Ma trận MTMM đơn giản chính là ma trận hệ số tương quan giữa các biến chỉ báo của các cấu trúc/biến tiềm ẩn. Trước khi đi vào công thức tính toán chỉ số HTMT, chúng ta sẽ hình dung qua ma trận MTMM của hai cấu trúc Y_1 và Y_2 được mô tả ở **Hình 31** để hiểu được các tính chỉ số HTMT.



Hình 31. Ma trận MTMM (Hair và cộng sự, 2022: trang 121).

Theo Hair và cộng sự (2022: trang 122), tỷ số HTMT đơn giản là tỷ số của các hệ số tương quan chéo (**between-trait correlations**) trên các hệ số tương quan nội bộ (**within-trait correlations**). Cụ thể hơn, tỷ số HTMT là tỷ số giá trị trung bình của tất cả các hệ số tương quan của các biến chỉ số đo lường các cấu trúc/biến tiềm ẩn khác nhau [**tức trung bình của tất cả các hệ số tương quan trong hình chữ nhật**] so với giá trị trung bình hình học của các hệ số tương quan trung bình của các biến chỉ số đo lường cùng cấu trúc/biến tiềm ẩn [**tức căn bậc hai của tích các giá trị trung bình các hệ số tương trong mỗi hình tam giác**]. Để đảm bảo tính chất khác biệt giữa các cấu trúc/biến tiềm ẩn thì các hệ số trong hình chữ nhật càng nhỏ càng tốt; điều này có nghĩa ít có sự tương quan giữa các biến chỉ số thuộc các cấu trúc/biến tiềm ẩn khác nhau. Công thức này được tính như sau:

$$HTMT(Y_1Y_2) = \frac{mean(R_{Y_1Y_2})}{\sqrt{mean(R_{Y_1Y_1}).mean(R_{Y_2Y_2})}} \quad (18)$$

Trong đó, $R_{Y_1Y_2}$ là ma trận các hệ số tương quan giữa mỗi biến chỉ báo của Y_1 và Y_2 (tức là các hệ số tương quan trong hình chữ nhật), và $R_{Y_1Y_1}$ ($R_{Y_2Y_2}$) là ma trận các hệ số tương quan giữa mỗi biến chỉ báo của Y_1 (Y_2) (Franke & Sarstedt, 2019).

Với các hệ số tương quan (chéo và nội bộ) trong **Hình 31**, ta dễ dàng tính được các giá trị trung bình trong công thức (18) như sau:

$$\text{Mean}(R_{Y_1Y_2}) = (0.426 + 0.339 + 0.393 + 0.423 + 0.345 + 0.385 + 0.274 + 0.235 + 0.250)/9 = \mathbf{0.341}$$

$$\text{Mean}(R_{Y_1Y_1}) = (0.770 + 0.701 + 0.665)/3 = \mathbf{0.712}$$

$$\text{Mean}(R_{Y_2Y_2}) = (0.574 + 0.318 + 0.335)/3 = \mathbf{0.409}$$

Vì thế,

$$HTMT(Y_1Y_2) = \frac{0.341}{\sqrt{0.712 \cdot 0.409}} = 0.632$$

Về nguyên tắc, trung bình của các hệ số tương quan chéo Heterotrait-heteromethod giữa các biến chỉ báo của cả hai cấu trúc/biến tiềm ẩn Y_1 và Y_2 phải càng nhỏ hơn trung bình hình học các hệ số tương quan Monotrait-heteromethod của từng cấu trúc/biến tiềm ẩn Y_1 và Y_2 càng tốt. Khi đó, hai cấu trúc/biến tiềm ẩn Y_1 và Y_2 chủ yếu chia sẻ sự biến thiên cho các biến chỉ báo trong nội bộ của mỗi cấu trúc/biến tiềm ẩn mà ít chia sẻ sự biến thiên cho nhau. Lúc này, tập các biến chỉ báo của cấu trúc/biến tiềm ẩn Y_1 sẽ đảm bảo tính phân biệt với tập các biến chỉ báo của cấu trúc/biến tiềm ẩn Y_2 .

Henseler và cộng sự (2015) đưa ra đề xuất về hai **ngưỡng đánh giá giá trị phân biệt** giữa tập các biến chỉ báo của các cấu trúc/biến tiềm ẩn A và của cấu trúc/biến tiềm ẩn B như sau:

- Nếu tỷ số $HTMT(A,B) > 0.9$, khó đạt được giá trị phân biệt giữa hai cấu trúc/biến tiềm ẩn A và B . Có nghĩa là dữ liệu của tập các biến chỉ báo của A và B là khá tương đồng với nhau.
- Nếu tỷ số $HTMT(A,B) \leq 0.85$, đạt được giá trị phân biệt giữa hai cấu trúc/biến tiềm ẩn A và B .

Lưu ý rằng, tỷ số HTMT chỉ đánh giá tính phân biệt cho từng cặp cấu trúc/biến tiềm ẩn trong mô hình đo lường kết quả. Hơn nữa, giữa giá trị hội tụ và giá trị phân biệt có xu hướng đồng hành với nhau; nghĩa là tính hội tụ (trong từng cấu trúc/biến tiềm ẩn) càng cao thì tính phân biệt (giữa các cặp cấu trúc/biến tiềm ẩn) càng được đảm bảo.

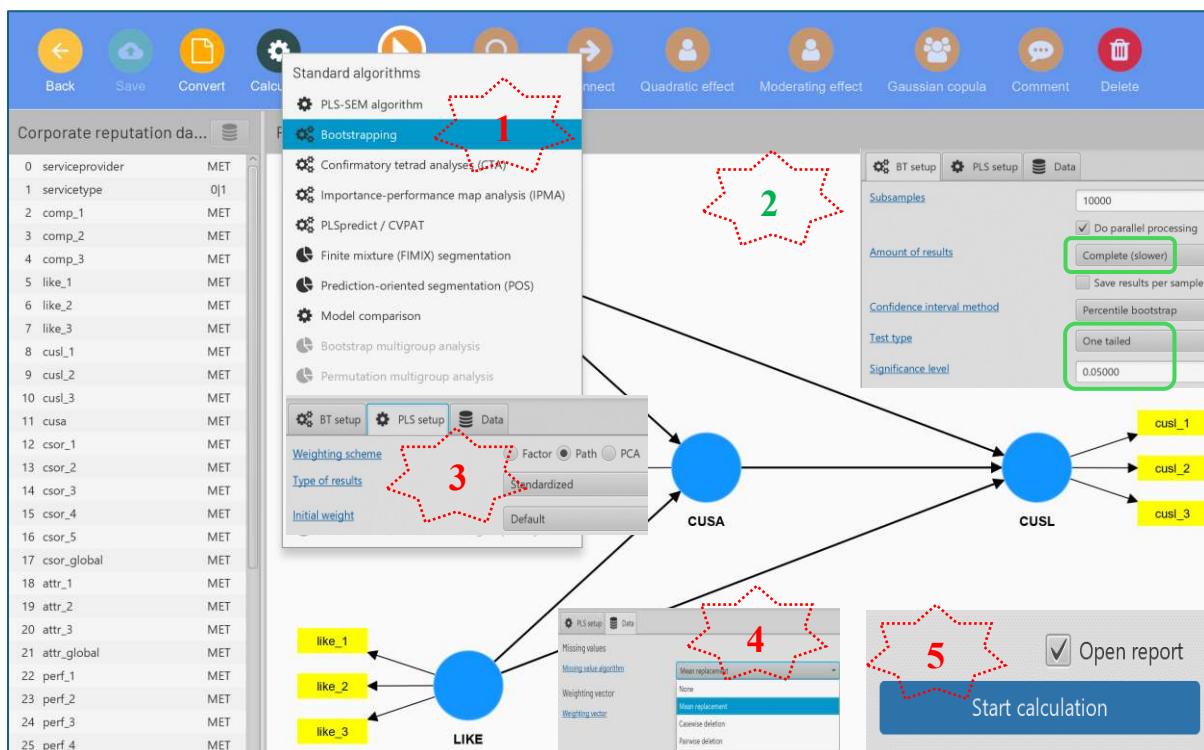
Trên SmartPLS 4 thì giá trị phân biệt được xác định bằng cách **Quality criteria → Discriminant validity**, chọn dạng ma trận, bảng hay đồ thị như **Hình 32**.



Hình 32. Giá trị phân biệt trong mô hình đơn giản.

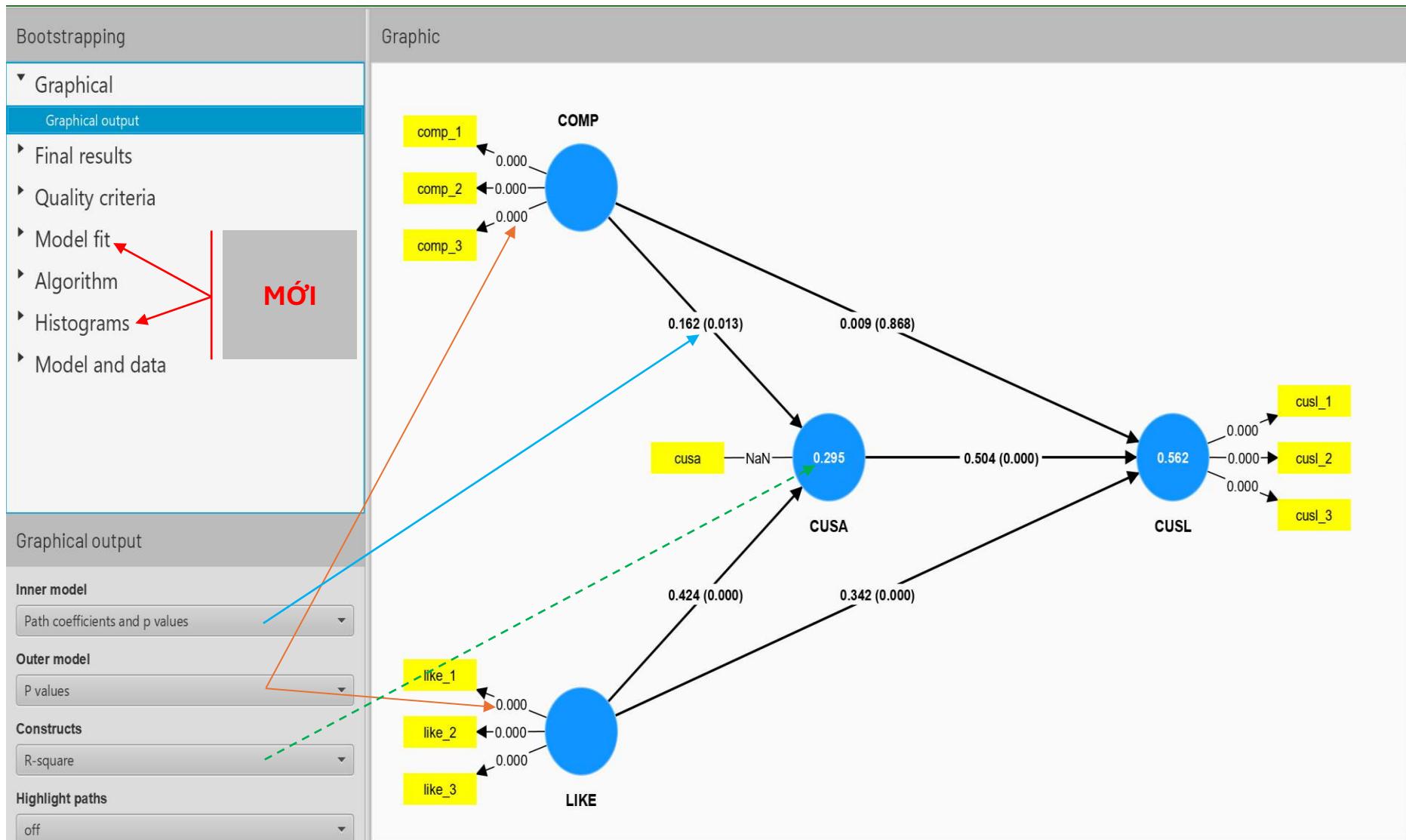
Theo Hair và cộng sự (2022: trang 123-24), thì mô hình PLS-SEM không dựa trên bất kỳ các giả định nào về phân phối xác suất, nên các kiểm định mức ý nghĩa thông thường không thể được áp dụng để kiểm định xem liệu thống kê HTMT từ kết quả ước lượng (như **Hình 32**) có khác ngưỡng đánh giá giá trị phân biệt (< 0.85) một cách có ý nghĩa thống kê hay không. Ví dụ, tỷ số HTMT_(LIKE,COMP) = 0.78 có chắc chắn nhỏ hơn 0.85 một cách có ý nghĩa thống kê ở 5% hay không? Để trả lời câu hỏi này, ta cần xây dựng khoảng tin cậy cho thống kê HTMT. Kỹ thuật mô phỏng hay thuật toán Bootstrapping là cách thường dùng trong các mô hình PLS-SEM.

Trước khi thực hiện chạy mô phỏng (**Bootstrapping**), bạn nên kiểm tra lại mình đã lưu mô hình đã được thực hiện chưa. Nếu chưa thì bấm “SAVE” để lưu lại. Trở lại với hướng dẫn ở **Hình 22**, khi chọn “**Calculate**” thay vì “**PLS-SEM algorithm**” bây giờ ta chọn “**Bootstrapping**”; vào “**BT setup**” nhập số lần thực hiện mô phỏng (thường là 10.000 lần), mức ý nghĩa (thường là 5%,), loại kiểm định (thường là hai phía), và đặc biệt quan trọng là với “**Amount of results**” phải chọn “**Complete (slower)**” thì mới có các kết quả liên quan đến “**Quality criteria**”, bao gồm tỷ số HTMT. Các bước còn lại sẽ tương tự như hướng dẫn ở **Hình 22**. Như vậy kỹ thuật mô phỏng này bao gồm năm bước (sau khi đã tạo mô hình PLS-SEM) như **Hình 33**.



Hình 33. Ước lượng mô phỏng mô hình đơn giản trên SmartPLS 4.

Sau khi chọn “**Start calculation**” thì ta được kết quả ước lượng như ở **Hình 34**. So với kết quả ước lượng thông thường như ở **Hình 23**, thì kết quả ước lượng theo kỹ thuật mô phỏng có một số đặc điểm vượt trội như ta có thể xác định được giá trị xác xuất, thống kê kiểm định t , khoảng tin cậy, đồ thị phân phối của các hệ số ước lượng p, l, w ; v.v. Bây giờ, trở lại ví dụ của ta về tỷ số **HTMT** như ở **Bảng 4**.



Hình 34. Kết quả ước lượng mô hình đơn giản theo thuật toán Bootstrapping trên SmartPLS 4.

Bảng 4. Tỷ số HTMT từ kết quả chạy mô phỏng.

	Original sample (O)	Sample mean (M)	5%	95%
CUSA <---> COMP	0.465	0.466	0.369	0.557
CUSL <---> COMP	0.532	0.533	0.434	0.624
CUSL <---> CUSA	0.755	0.755	0.698	0.808
LIKE <---> COMP	0.78	0.78	0.71	0.845
LIKE <---> CUSA	0.577	0.577	0.506	0.645
LIKE <---> CUSL	0.737	0.738	0.668	0.803

Kết quả chạy mô phỏng cho thấy các giá trị cao nhất trong khoảng tin cậy (5%-95%) đều thấp hơn ngưỡng giá trị (0.85). Điều này chứng tỏ các cấu trúc/biến tiềm ẩn trong mô hình PLS-SEM đơn giản đảm bảo giá trị phân biệt.

Sau cùng, chúng ta nên lập một bảng tổng hợp kết quả đánh giá mô hình đo lường kết quả, như gợi ý trong **Bảng 5**.

Bảng 5. Tóm tắt kết quả chính của mô hình đo lường kết quả.

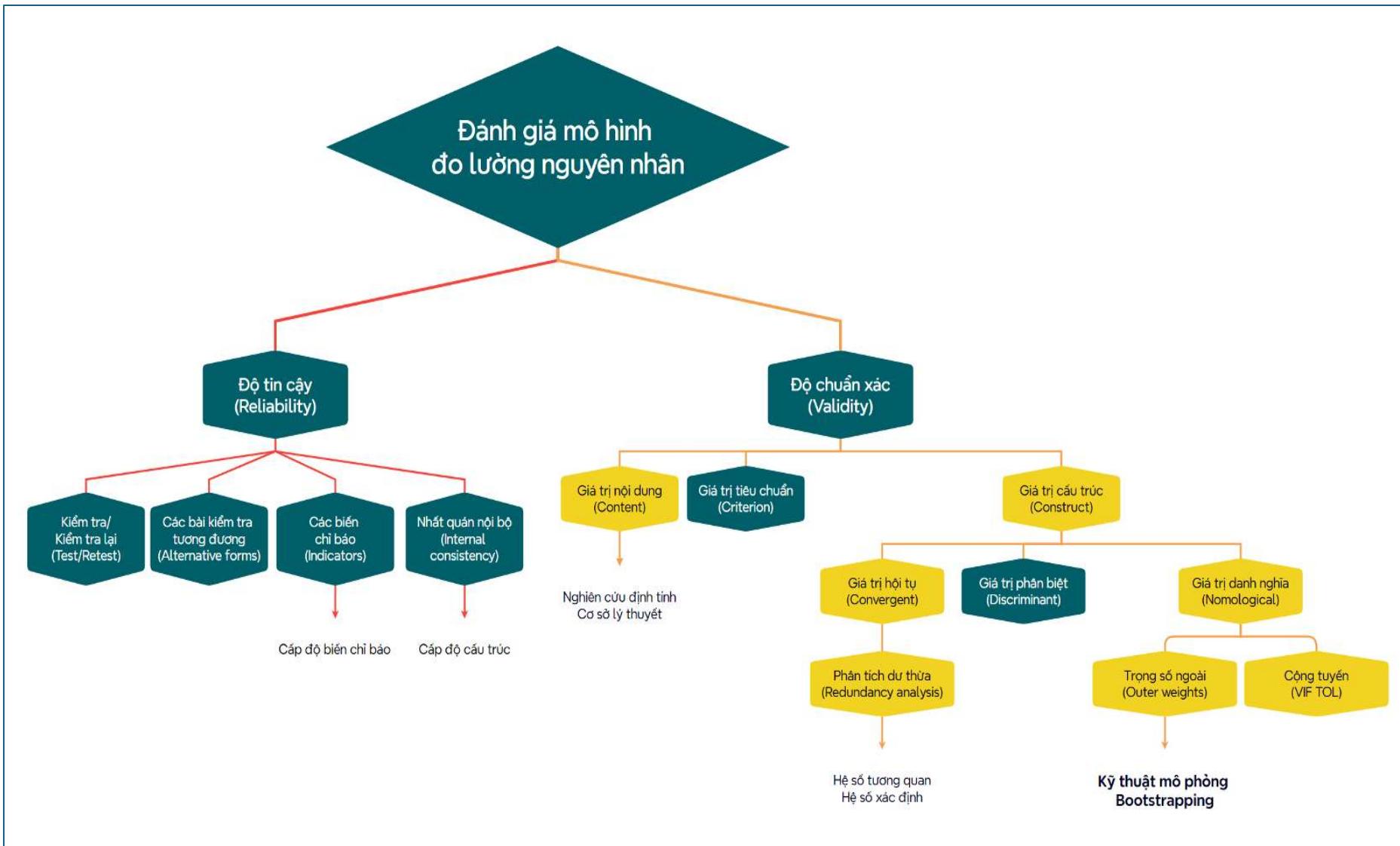
Latent Variable	Indicators	Convergent Validity			Internal Consistency Reliability			Discriminant Validity HTMT
		Loadings	Indicator Reliability	AVE	Cronbach's Alpha	Reliability ρ_A	Composite Reliability ρ_C	
		>0.70	>0.50	>0.50	0.60–0.90	0.60–0.90	0.60–0.90	
<i>COMP</i>	<i>comp_1</i>	0.858	0.736	0.681	0.776	0.832	0.865	Significantly lower than 0.85 (0.90)? Yes
	<i>comp_2</i>	0.798	0.637					
	<i>comp_3</i>	0.818	0.669					
<i>CUSL</i>	<i>cusl_1</i>	0.833	0.694	0.748	0.831	0.839	0.899	Yes
	<i>cusl_2</i>	0.917	0.841					
	<i>cusl_3</i>	0.843	0.711					
<i>LIKE</i>	<i>like_1</i>	0.879	0.773	0.747	0.831	0.836	0.899	Yes
	<i>like_2</i>	0.870	0.757					
	<i>like_3</i>	0.843	0.711					

Nguồn: Hair và cộng sự (2022: trang 135).

6.3 Đánh giá mô hình đo lường nguyên nhân

Theo Hair và cộng sự (2022: trang 141), các tiêu chí đánh giá thống kê cho các thang đo đo lường kết quả (reflective measurement scales¹⁰) không thể được áp dụng trực tiếp cho các mô hình đo lường nguyên nhân (formative measurement models) ở đó các biến

¹⁰ Thang đo (scales) ở đây có nghĩa là các cấu trúc/các biến đo lường.

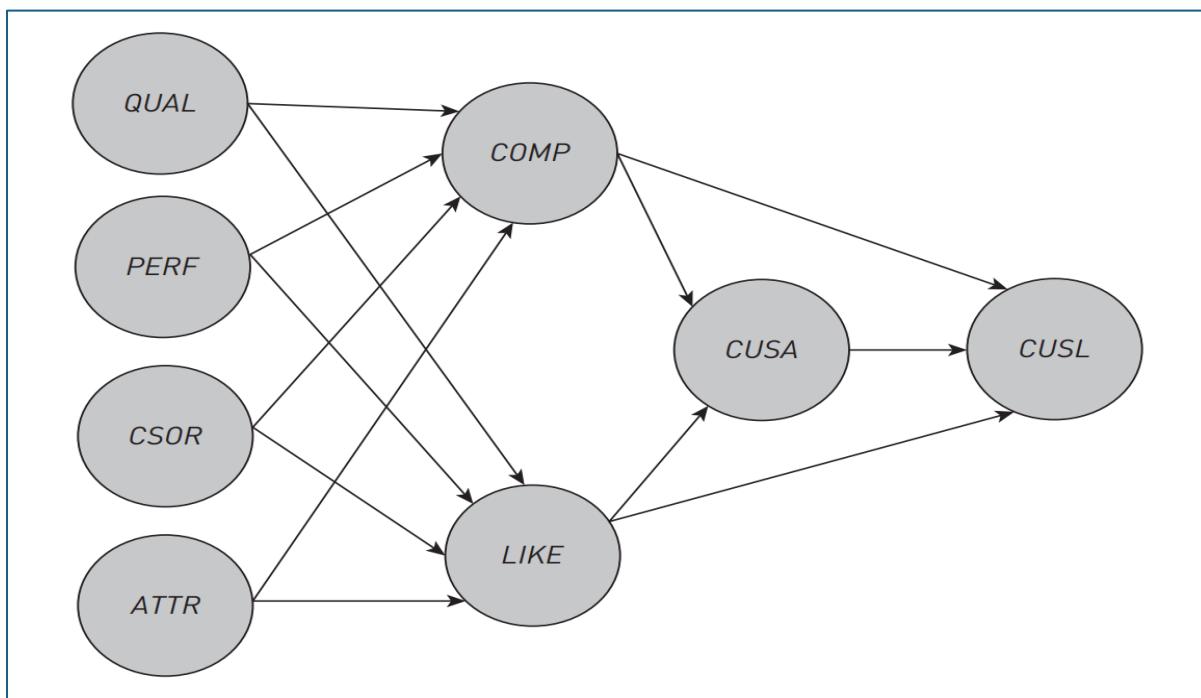


Hình 35. Nội dung đánh giá mô hình đo lường nguyên nhân.

chỉ báo có thể thể hiện các “nguyên nhân (**causes**)” độc lập của cấu trúc và vì thế không nhất thiết tương quan cao. Hơn nữa, các biến chỉ báo nguyên nhân (formative indicators) được giả định là **không có sai số**¹¹ (Bollen & Diamantopoulos, 2017), điều này có nghĩa là khái niệm độ tin cậy nhất quán nội bộ không còn thích hợp.

Đánh giá giá trị hội tụ và giá trị phân biệt của các cấu trúc đo lường nguyên nhân bằng cách sử dụng các tiêu chí tương tự các cấu trúc của các mô hình đo lường kết quả là không có ý nghĩa (Chin, 1998). Thay vào đó, các nhà nghiên cứu nên tập trung vào việc thiết lập giá trị nội dung (**content validity**) trước khi đánh giá các cấu trúc đo lường nguyên nhân. Bước này yêu cầu đảm bảo rằng các biến chỉ báo nguyên nhân phải bao hàm tất cả các khía cạnh (hoặc ít nhất là các khía cạnh quan trọng) của cấu trúc¹². Nội dung chính của đánh giá mô hình đo lường nguyên nhân được trình bày ở **Hình 35**.

Trước khi tiến hành các bước đánh giá mô hình đo lường nguyên nhân, cũng như mô hình cấu trúc ở phần tiếp theo, chúng ta hãy **mở rộng mô hình đơn giản** ở trên bằng cách đưa thêm các cấu trúc/biến tiềm ẩn nguyên nhân vào mô hình. Các cấu trúc/biến tiềm ẩn nguyên nhân được đưa vào mô hình mở rộng gồm: **QUAL, PERF, CSOR, và ATTR**; các biến này đã được giới thiệu và mô tả ở **Mục 3**. Mô hình mới được tạo ra bao gồm tất cả các cấu trúc/biến tiềm ẩn kết quả (COMP, LIKE, CUSL) VÀ CUSA như ở các phần trước và các cấu trúc/biến tiềm ẩn nguyên nhân (QUAL, PERF, CSOR, ATTR). Ta gọi đây là mô hình PLS-SEM mở rộng (**Extended model**).



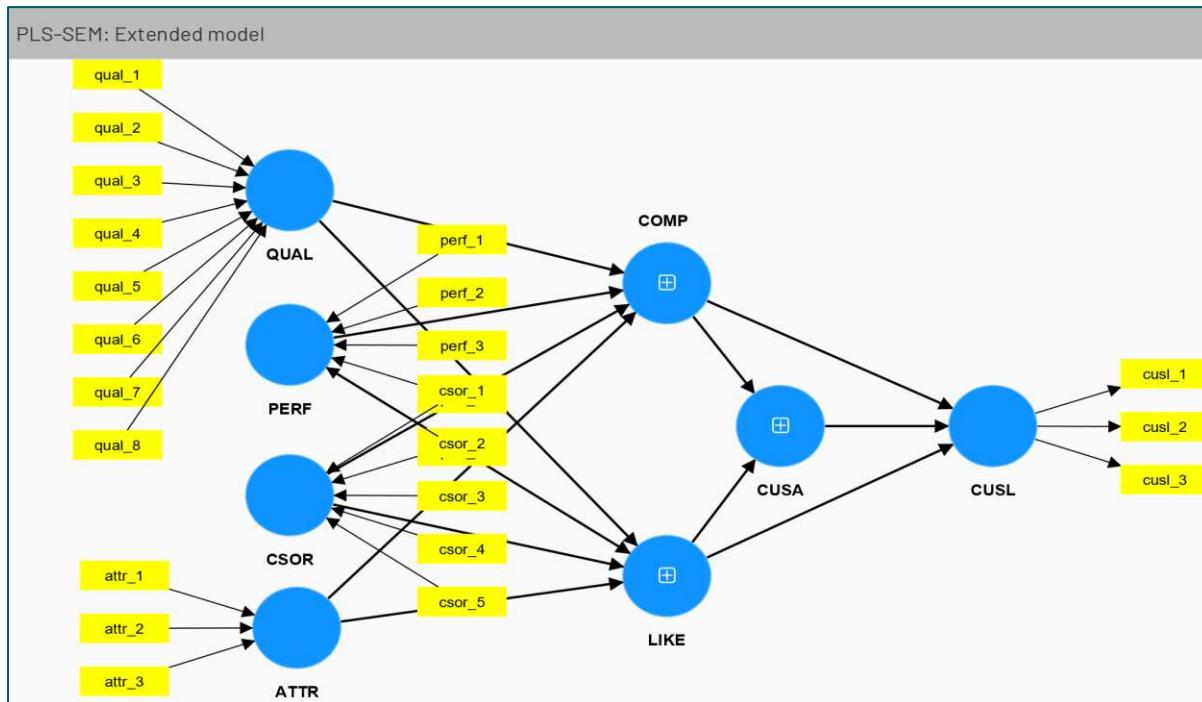
Hình 36. Mô hình mở rộng về lòng trung thành (Hair và cộng sự, 2022: trang 160).

¹¹ Các biến chỉ báo/biến quan sát nguyên nhân đóng vai trò là các biến giải thích cho cấu trúc/biến tiềm ẩn như **phương trình (2)**.

¹² Ý này liên quan đến khái niệm “**không có sai số**” ở trên. Nghĩa là người nghiên cứu cần phải lược khảo kĩ tài liệu, cơ sở lý thuyết liên quan và/hoặc nghiên cứu định tính để đảm bảo rằng các biến chỉ báo được chọn phải phản ánh đúng nhất bản chất/các khía cạnh khác nhau của khái niệm nghiên cứu (tức là các cấu trúc/biến tiềm ẩn).

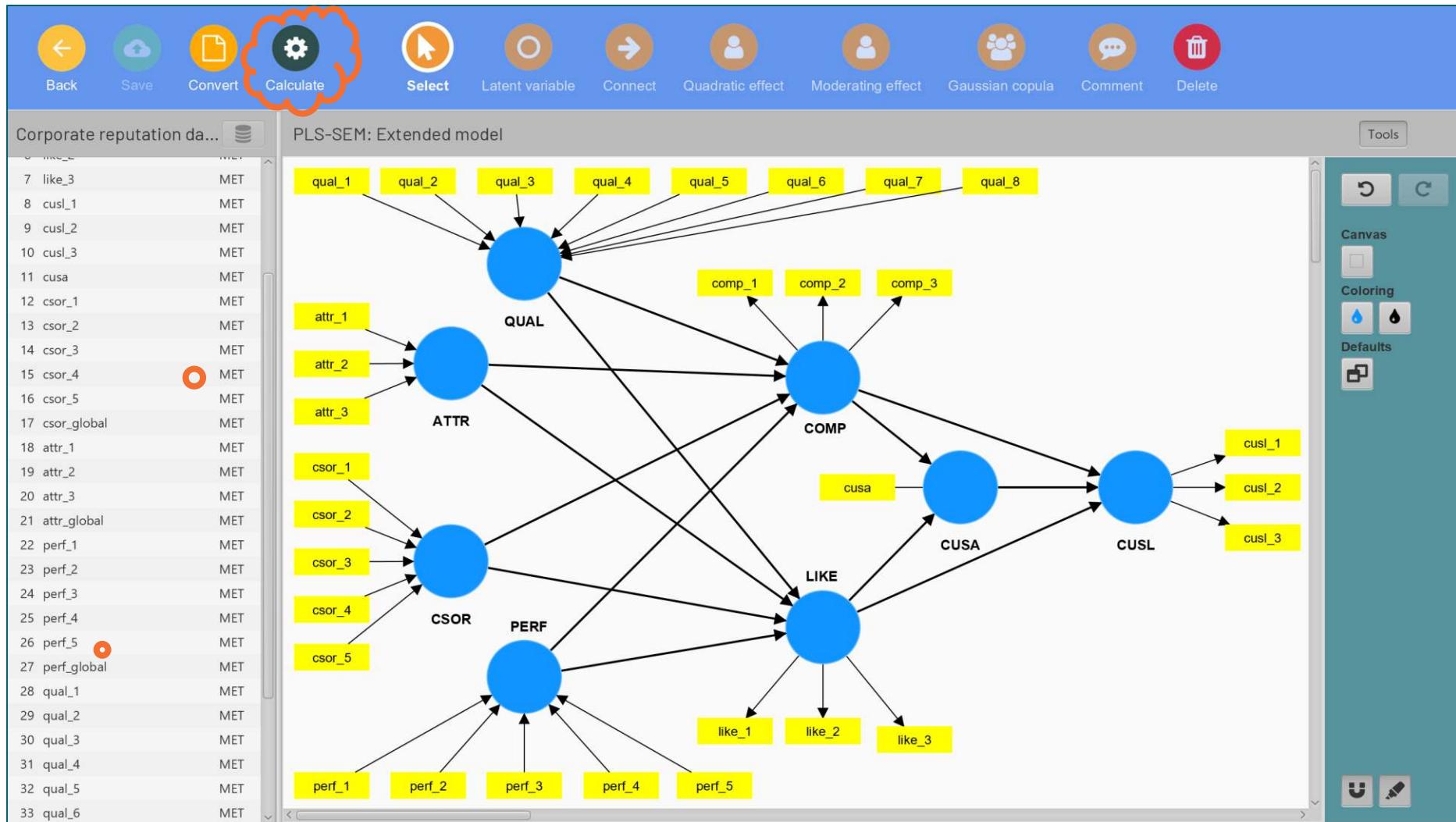
Để tạo mô hình PLS-SEM mở rộng như **Hình 36**, ta sẽ thực hiện các bước sau đây trên SmartPLS 4:

- Vào dự án **PLS-SEM2024** → Nhấp chuột phải vào **Simple model** → Chọn **Duplicate** → Đặt tên mô hình mới là **Extended model** → **Create**¹³.
- Nhấp đúp vào **Extended model** để mở → Chọn “**Select**” để kéo và chỉnh các cấu trúc sẵn có đến các vị trí thích hợp như ở **Hình 36** (Để dễ nhìn, nhấp chuột vào hình tròn của mỗi cấu trúc → nhấp chuột phải → chọn **Hide indicators**).
- Để chuột ở các vị trí thích hợp với các cấu trúc mới (QUAL, PERF, CSOR, ATTR), nhấp đúp để hiển thị “**Alpha**” và đặt tên cho từng cấu trúc mới; sau đó cân chỉnh cho các cấu trúc mới này thẳng hàng như **Hình 36**.
- Chọn “**Connect**”, vào kết nối các cấu trúc mới này với hai cấu trúc COMP và LIKE như **Hình 36**.
- Trong danh sách các biến chỉ báo ở hộp bên trái; chọn nhóm các biến chỉ báo thích hợp (như qual_1 qual_2 ... qual_8) bằng cách giữ Shift + phím mũi tên lên xuống; giữ và kéo bỏ qua hộp tròn của cấu trúc thích hợp (như QUAL), rồi nhả chuột. Làm tương tự cho các cấu trúc còn lại. Nhấp chuột phải vào các cấu trúc này, chọn “**Invert measurement model**”. Kết quả ban đầu như **Hình 37**.
- Cuối cùng, định dạng lại mô hình cho dễ nhìn, hiển thị các biến chỉ báo bị che ẩn bằng cách nhấp chuột vào hình tròn, nhấp chuột phải, rồi chọn **Show indicators**. Kết quả ta có sơ đồ hoàn chỉnh như **Hình 38**. Nhớ “**SAVE**”.



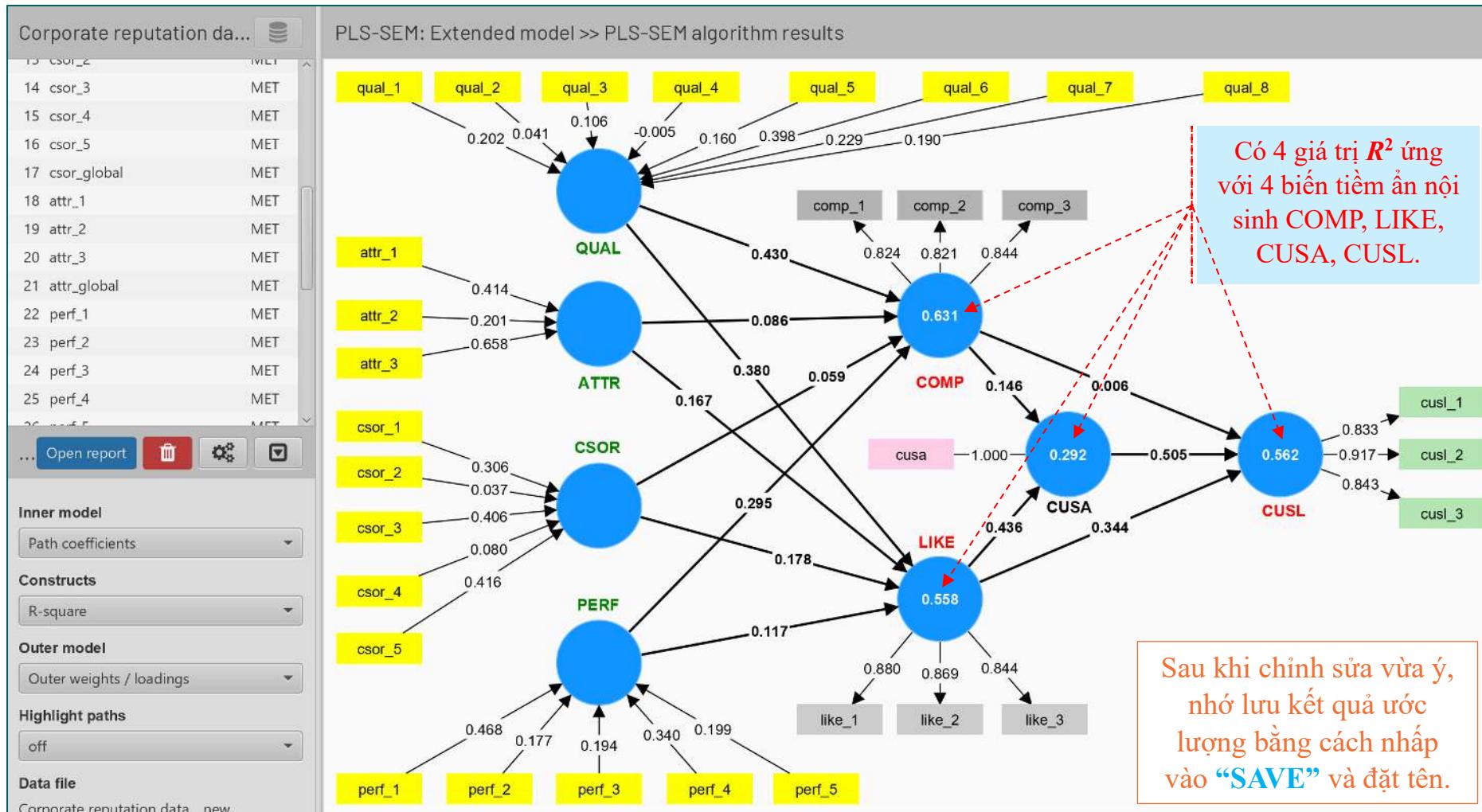
Hình 37. Tạo mô hình mở rộng về lòng trung thành của khách hàng.

¹³ Cách khác: Vào dự án **PLS-SEM2024** → Nhấp vào biểu tượng **PLS-SEM** → Đặt tên mô hình mới là **Extended model** → **Create**; Sau đó, nhấp chuột phải vào **Simple model**, chọn **Copy resource**; Trở lại **Extended model** vừa tạo, nhấp chuột phải, chọn **Paste resource**. Tuy nhiên, cách này sẽ mất nhiều thời gian hơn.



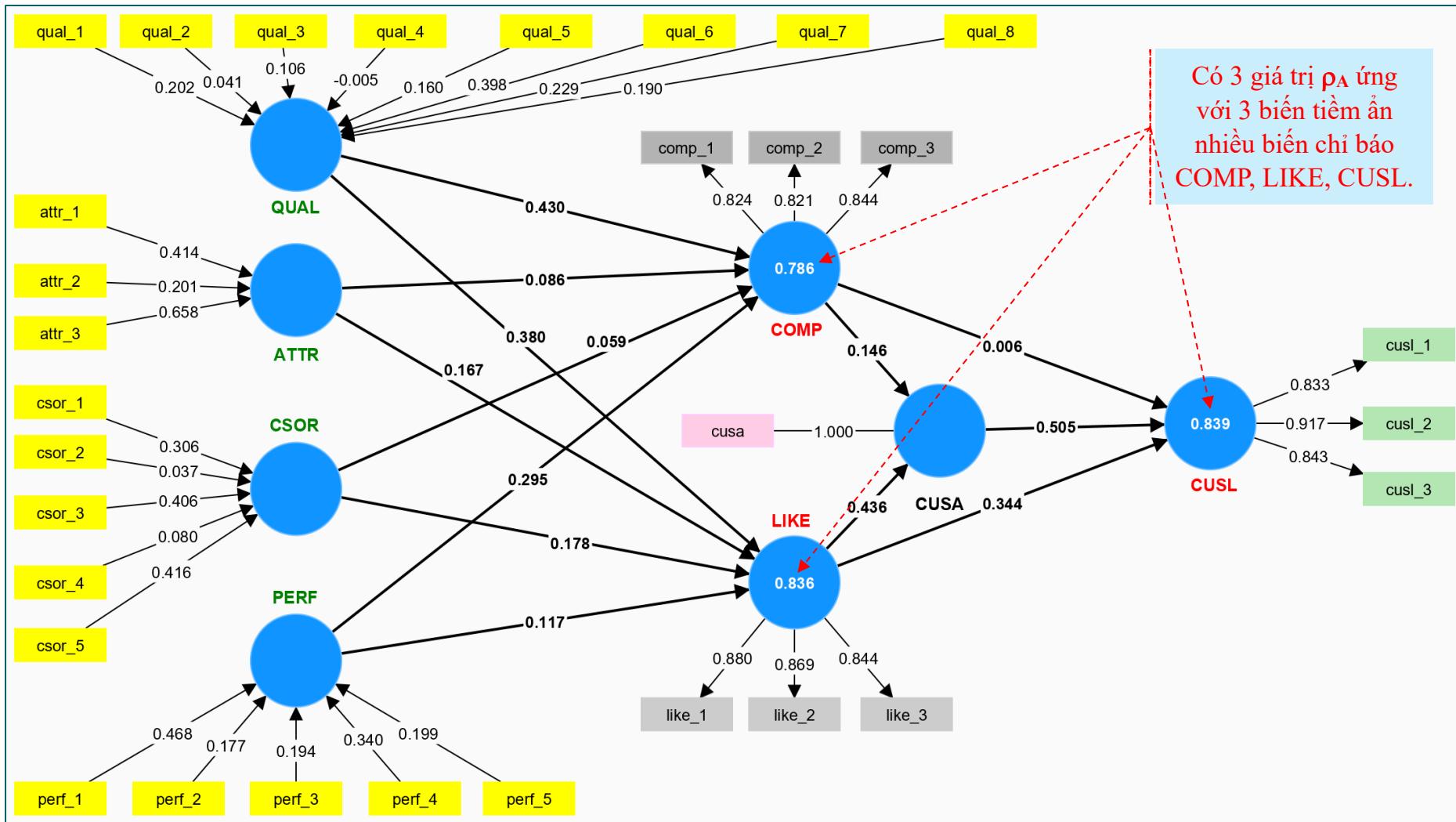
Hình 38. Hoàn chỉnh mô hình PLS-SEM mở rộng về lòng trung thành của khách hàng.

- **Ước lượng:** Chọn “Calculate” → **PLS-SEM algorithm** → Kiểm tra **PLS setup** và **Data** → **Start calculation**. Kết quả ở **Hình 39**.



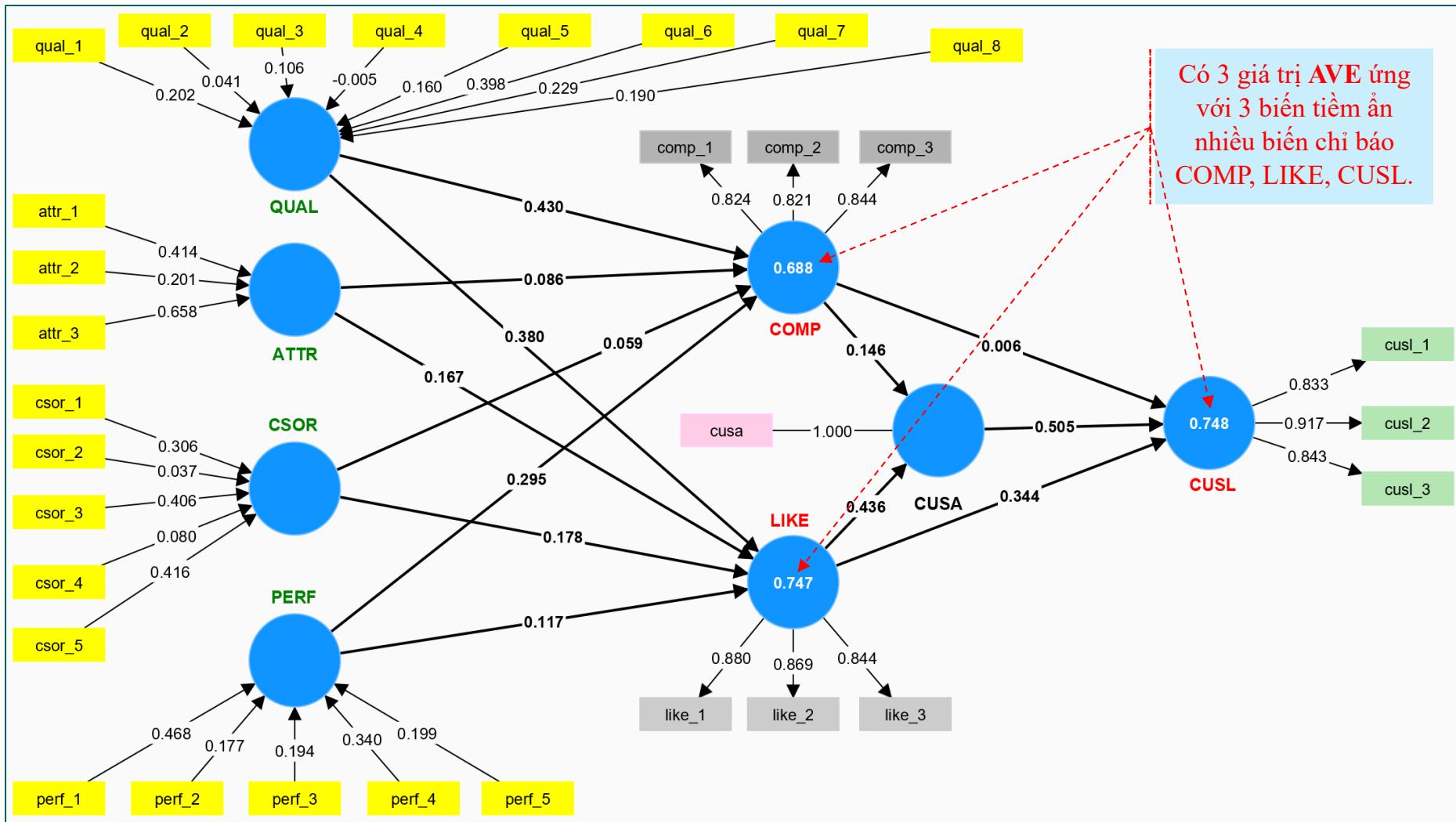
Hình 39. Kết quả ước lượng mô hình PLS-SEM mở rộng về lòng trung thành của khách hàng (Các hệ số l , p , w ; và R^2).

Lưu ý: Ta có thể chỉnh sửa vị trí các hình, tên cấu trúc, giá trị thống kê bằng cách giữ chuột và kéo đến vị trí vừa ý sao cho có các hiển thị rõ ràng, dễ nhìn nhất (tránh bị khuất, trùng lặp). Ngoài ra, ta cũng có thể đổi màu cho dễ nhìn (chọn và sử dụng hộp công cụ).



Hình 40. Kết quả ước lượng mô hình PLS-SEM mở rộng về lòng trung thành của khách hàng (Các hệ số l , p , w ; và ρ_A).

Lưu ý: Bạn hãy so sánh các hệ số l , p , và ρ_A của mô hình này với mô hình đơn giản?

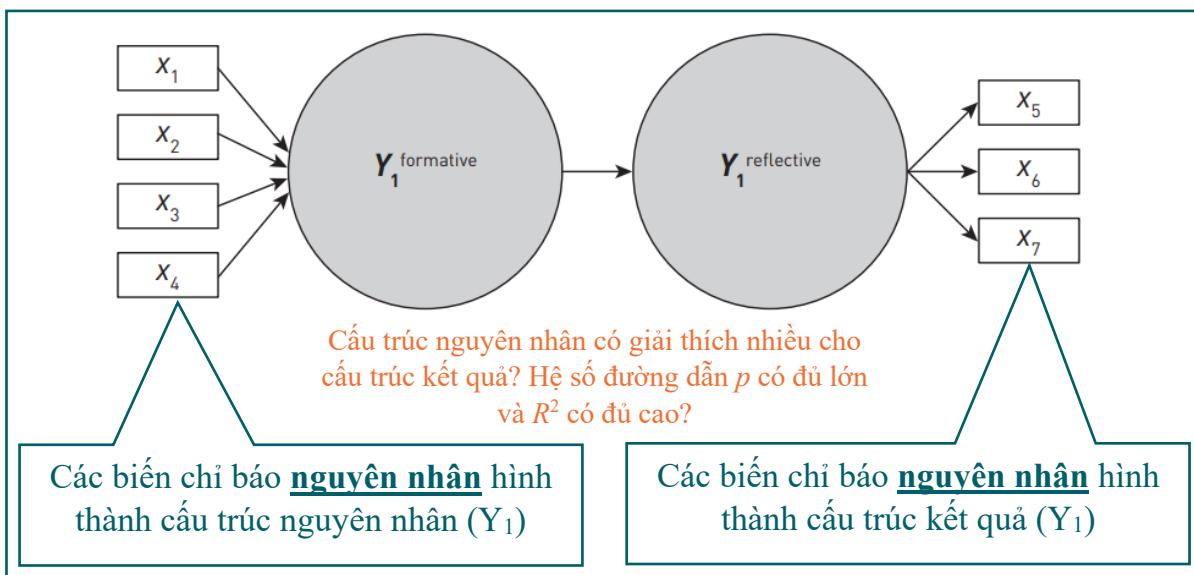


Hình 41. Kết quả ước lượng mô hình PLS-SEM mở rộng về lòng trung thành của khách hàng (Các hệ số l , p , w ; và AVE).

Lưu ý: Bạn hãy so sánh các hệ số l , p , và AVE của mô hình này với mô hình đơn giản? Hãy tự thực hành với các giá trị thống kê khác!

Bước 1: Đánh giá giá trị phân biệt (Discriminant validity).

Theo Hair và cộng sự (2022: trang 143), khi đánh giá các mô hình đo lường nguyên nhân, ta phải kiểm định liệu cấu trúc được đo bằng các biến chỉ báo nguyên nhân có tương quan cao với một thước đo kết quả của cùng một cấu trúc. Loại kiểm định này được gọi là **phân tích dư thừa** (redundancy analysis) (Chin, 1998). Thuật ngữ “phân tích dư thừa” bắt nguồn từ việc thông tin trong mô hình trở nên dư thừa theo ý nghĩa là nó được bao gồm trong cấu trúc nguyên nhân và lại được bao gồm trong cấu trúc kết quả. Cụ thể, ta phải sử dụng các trúc đo lường nguyên nhân như một biến tiềm ẩn ngoại sinh để dự đoán một biến tiềm ẩn nội sinh được hình thành từ một hay nhiều biến chỉ báo kết quả như ở **Hình 42**.



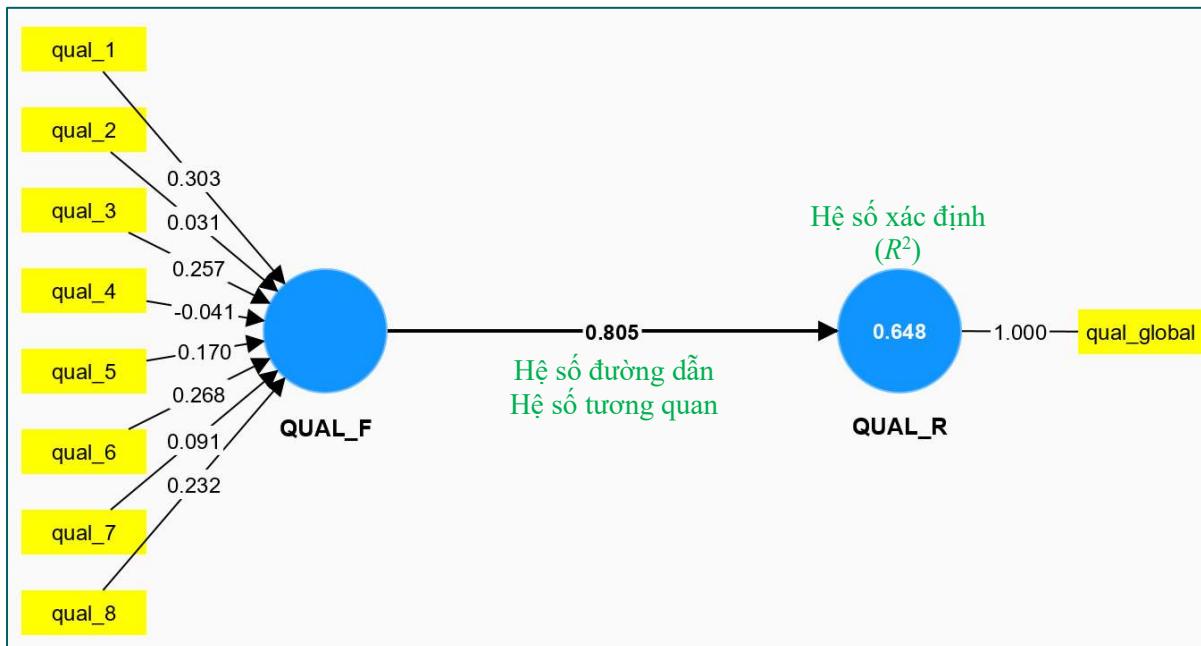
Hình 42. Phân tích dư thừa để đánh giá giá trị hội tụ.

Các biến chỉ báo X_1, X_2, \dots, X_7 là cùng thể loại; tức chúng cùng được thu thập để đo lường cấu trúc Y_1 . Nếu làm như vậy thì sẽ gây khó khăn lúc thực hiện khảo sát hoặc chất lượng các câu trả lời của những người được phỏng vấn bởi vì cùng lúc họ phải trả lời quá nhiều câu hỏi về một khái niệm quan tâm. Một cách thực hành phổ biến là chúng ta chỉ dùng một câu hỏi (một biến chỉ báo) chung thay cho các biến X_5, X_6, X_7 .

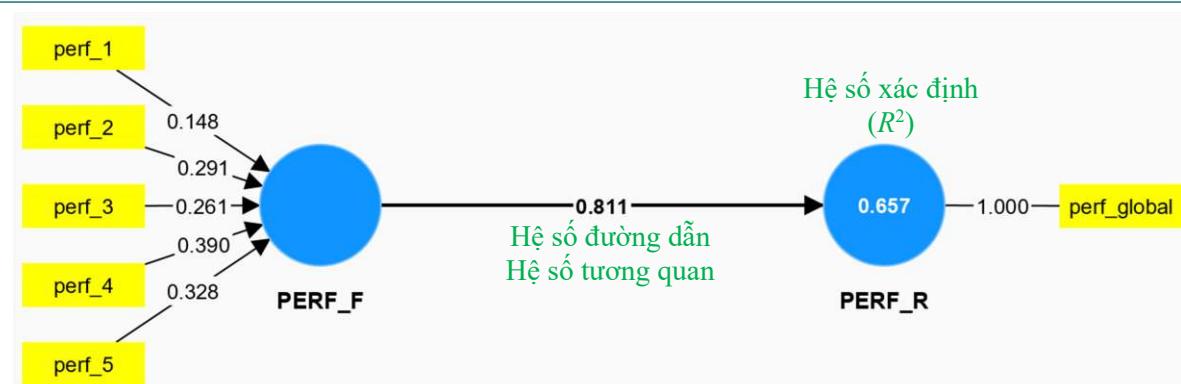
Hair và cộng sự (2022: trang 143) cho rằng độ lớn của hệ số đường dẫn kết nối hai biến tiềm ẩn (của cùng một khái niệm/cấu trúc) chỉ cho ta biết độ chuẩn xác (validity) của tập các biến chỉ báo nhằm đo lường cấu trúc quan tâm. Lý tưởng nhất là độ lớn của hệ số đường dẫn (p) bằng **0.8** trở lên; nhưng với giá trị từ **0.7** trở lên là cũng thỏa đáng cho đường dẫn giữa $Y_1^{\text{nguyên nhân}}$ và $Y_1^{\text{kết quả}}$. Với các giá trị này thì khi chuyển sang các giá trị R^2 tương đương khoảng **0.64** hay ít nhất là **0.5**. Nếu phân tích chỉ ra là thiếu giá trị hội tụ (tức là giá trị $R^2 < 0.5$), thì các biến chỉ báo nguyên nhân của cấu trúc $Y_1^{\text{nguyên nhân}}$ không đóng góp mức độ đủ cho nội dung cần được nhắm đến. Nếu điều đó xảy ra thì về mặt lý thuyết ta cần phải chỉnh sửa bằng cách đưa thêm các biến chỉ báo nguyên nhân khác vào bảng câu hỏi.

Để thực hiện phân tích dư thừa trên SmartPLS 4 (và có kết quả như **Exhibit 5.21**, Hair và cộng sự, 2022: trang 172), ta cần phải tạo mới các mô hình cấu trúc đơn giản này theo các bước sau đây.

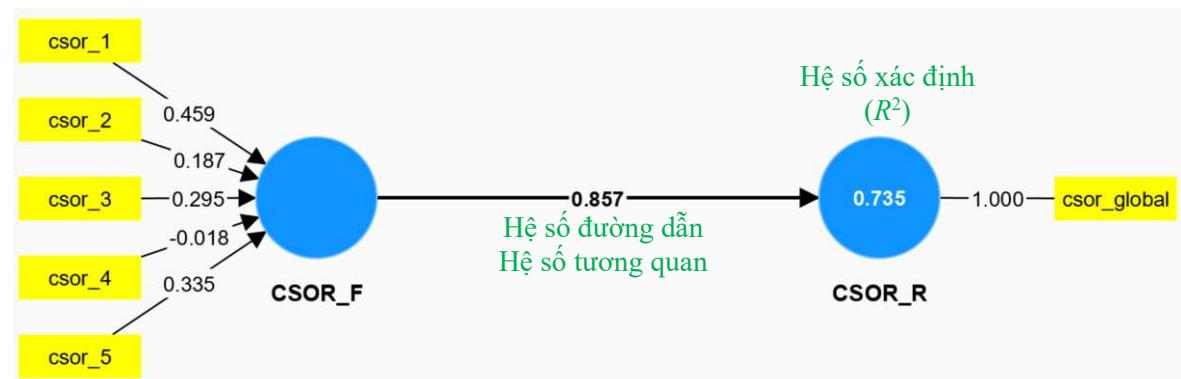
- Vào dự án **PLS-SEM2024** → Nhập vào biểu tượng **PLS-SEM** → Đặt tên mô hình mới là **Redundancy analysis _ QUAL** → **Create**.
- Nhấp đúp vào **Redundancy analysis _ QUAL** để mở → Chọn “**Select**” → Đặt chuột ở các vị trí thích hợp, nhấp đúp để hiển thị “**Alpha**” và đặt tên cấu trúc **QUAL_F** (tức là $QUAL^{\text{formative}}$); Nhấp đúp ở vị trí thích hợp khác để hiển thị “**Alpha**” và đặt tên cấu trúc **QUAL_R** (tức là $QUAL^{\text{reflective}}$)
- Chọn “**Connect**”, vào kết nối hai cấu trúc mới này với nhau.
- Trong danh sách các biến chỉ báo ở hộp bên trái; chọn nhóm các biến chỉ báo thích hợp (**qual_1 qual_2 ... qual_8**) bằng cách giữ Shift + phím mũi tên lên xuống; giữ và kéo bỏ qua hộp tròn của cấu trúc **QUAL_F**, rồi nhả chuột; Nhấp chuột phải vào cấu trúc này, chọn “**Invert measurement model**”.
- Tương tự, chọn biến chỉ báo **qual_global**, giữ và kéo bỏ qua hộp tròn của cấu trúc **QUAL_R**, rồi nhả chuột. Lưu mô hình bằng cách nhập “**SAVE**”.
- Làm tương tự cho các cấu trúc còn lại (**PERF_F → PERF_R; CSOR_F → CSOR_R; ATTR_F → ATTR_R**). Cách làm nhanh nhất là sử dụng **Duplicate** (sau đó đổi tên cấu trúc; chọn và xóa các biến chỉ báo cũ; xác định lại các biến chỉ báo mới). Tổng cộng là 4 mô hình mới ứng với 4 cấu trúc nguyên nhân.
- Mở từng mô hình vừa tạo ra, thực hiện ước lượng với PLS-SEM algorithm. Cho hiển thị **R²** ở “**Constructs**” và **Path coefficients** ở “**Inner model**”. Lưu ý: với mô hình hồi quy đơn giữa hai biến tiềm ẩn thì hệ số đường dẫn **p** cũng chính là hệ số tương quan giữa hai biến. Bạn có thể kiểm tra bằng cách đổi thành **Correlations** ở “**Inner model**”. Ta có kết quả như **Hình 43a-d**.



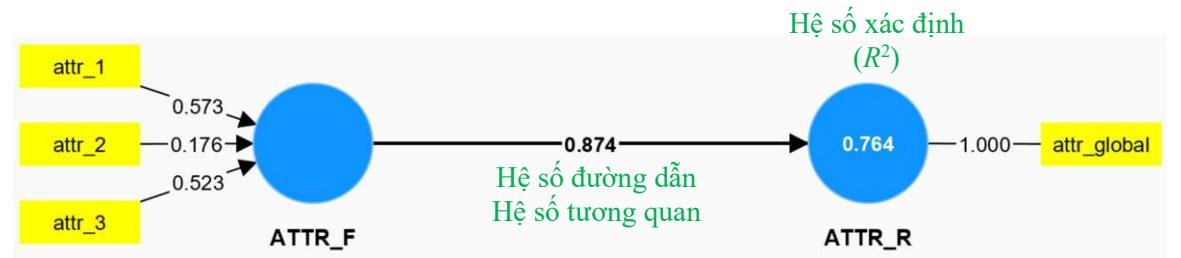
Hình 43a. Phân tích dư thừa cho cấu trúc QUAL.



Hình 43b. Phân tích dư thừa cho cấu trúc **PERF**.



Hình 43c. Phân tích dư thừa cho cấu trúc **CSOR**.



Hình 43d. Phân tích dư thừa cho cấu trúc **ATTR**.

Các kết quả phân tích dư thừa ở các **Hình 43a-d** chỉ ra rằng độ lớn các hệ số đường dẫn đều cao hơn giá trị ngưỡng đề xuất (> 0.8), và giá trị hệ số xác định R^2 cũng cao hơn giá trị ngưỡng đề xuất (> 0.64). Kết quả này giúp ta khẳng định rằng tất cả các cấu trúc nguyên nhân đều đáp ứng tiêu chí giá trị hội tụ.

Bước 2: Đánh giá giá trị phân biệt (Discriminant validity).

Trong mô hình đo lường nguyên nhân, chúng ta không kỳ vọng các biến chỉ báo có tương quan cao với nhau như trong mô hình đo lường kết quả. Khi hai biến chỉ báo nguyên nhân có hệ số tương quan cao hoặc rất cao thì ta gọi là vấn đề cộng tuyến (**collinearity**). Mở rộng ra, khi có nhiều biến chỉ báo nguyên nhân tương quan cao với nhau thì ta gọi đó là hiện tượng đa cộng tuyến (**multicollinearity**).

Vấn đề nghiêm trọng nhất là khi hai hay nhiều biến chỉ báo nguyên nhân của cùng một tập (tức cùng phản ánh một khái niệm/cấu trúc) có thông tin hoàn toàn giống nhau (gọi là sự cộng tuyến hay đa cộng tuyến hoàn hảo). Khả năng này có thể xảy ra bởi vì cùng một biến chỉ báo được nhập liệu hai lần hoặc bởi vì một biến chỉ báo là một kết hợp tuyến tính của một biến chỉ báo khác (như doanh số tính bằng đơn vị khối lượng và doanh số tính theo giá trị). *Khi gặp phải những trường hợp này, PLS-SEM không thể ước lượng được các hệ số*; cụ thể, không thể ước lượng các trọng số ngoài (w). Sự cộng tuyến hay đa cộng tuyến cũng có thể xảy ra ở mô hình cấu trúc nếu như các biến chỉ báo dư thừa (redundant indicators) được sử dụng như các hạng mục đơn để đo lường hai hay nhiều cấu trúc.

Tại sao sự cộng tuyến hay đa cộng tuyến lại quan trọng? Theo [Hair và cộng sự \(2022: trang 145\)](#) thì các mức độ cộng tuyến hay đa cộng tuyến cao giữa các biến chỉ báo nguyên nhân là một vấn đề rất quan trọng bởi vì chúng có tác động lên sự ước lượng các trọng số [xem **phương trình (2)**] và ý nghĩa thống kê của các trọng số. Hậu quả của (đa) cộng tuyến cao nằm ở hai khía cạnh. Một, (đa) cộng tuyến làm gia tăng các sai số chuẩn của ước lượng, và vì thế gia tăng khả năng giá trị ước lượng của các trọng số w không có ý nghĩa thống kê. Hai, (đa) cộng tuyến cao có thể dẫn đến việc các trọng số w bị ước lượng sai (ta hay gọi là ước lượng chêch – **biased estimates**) và sai dấu. Để hiểu rõ hơn về vấn đề này, bạn hãy tham khảo một số giáo trình về kinh tế lượng¹⁴.

Để đánh giá mức độ (đa) cộng tuyến, các nhà nghiên cứu thường tính hệ số độ chấp nhận **TOL** (tolerance) hoặc nhân tố khuếch đại phương sai **VIF** (variance inflation factor). TOL thể hiện lượng phương sai của một biến chỉ báo không được giải thích bởi các biến chỉ báo khác trong cùng cấu trúc. VIF là nghịch đảo của TOL. Cách tính giá trị TOL như sau:

- Hồi quy một biến chỉ báo (X_k) theo tất cả các biến chỉ báo còn lại của cùng một cấu trúc; tính giá trị R^2 .
- Tính TOL theo công thức sau:

$$TOL_{X_k} = 1 - R_{X_k}^2 \quad (19)$$

Như vậy, VIF của biến chỉ báo X_k được theo công thức sau:

$$VIF_{X_k} = \frac{1}{TOL_{X_k}} \quad (20)$$

Một quy tắc chung là các giá trị VIF *bằng hoặc lớn hơn 5* cho biết tồn tại (đa) cộng tuyến giữa các biến chỉ báo của các cấu trúc đo lường nguyên nhân. Tuy nhiên, [Becker và cộng sự \(2015\)](#), [Mason & Perreault \(1991\)](#) cho rằng các vấn đề (đa) cộng tuyến cũng có thể xảy ra ở các giá trị VIF bằng 3. Cho nên, [Hair và cộng sự \(2022: trang 147\)](#) khuyến nghị các giá trị VIF lý tưởng (để xác định không tồn tại vấn đề cộng tuyến) là

¹⁴ Ví dụ như: [Asteriou & Hall \(2021: chapter 5\)](#); [Gujarati & Porter \(2009, chapter 10\)](#); [Gujarati \(2014, chapter 4\)](#); [Hill và cộng sự \(2018: chapter 6\)](#); [Studenmund \(2017: chapter 8\)](#); v.v.

bằng hoặc thấp hơn 3. Điều này có nghĩa là giá trị TOL lý tưởng sẽ bằng hoặc thấp hơn **0.333**. Nếu mức độ (đa) cộng tuyến rất cao ($VIF \geq 5$), thì ta phải xem xét loại bỏ một trong số những biến chỉ báo tương ứng. Tuy nhiên, việc loại bỏ biến chỉ báo đó phải đảm bảo rằng các biến chỉ báo còn lại vẫn chứa đựng đầy đủ nội dung của cấu trúc (biến tiềm ẩn) liên quan. Một cách khác là kết hợp hai hay nhiều biến chỉ báo cộng tuyến cao với nhau thành một biến chỉ báo đơn tổng hợp (**single composite indicator**) cũng là một lựa chọn nên xem xét.

Trở lại với ví dụ về mô hình mở rộng của chúng ta. Để kiểm định (đa) cộng tuyến có tồn tại trong các mô hình đo lường nguyên nhân hay không, ta chọn: *Quality Criteria → Collinearity Statistics (VIF) → Outer VIF Values*. Kết quả (xin không trình bày chi tiết ở đây) cho thấy VIF cao nhất chỉ có giá trị 2.2 (biến chỉ báo *qual_5*). Điều này cho thấy tất cả các giá trị VIF đều thấp hơn giá trị ngưỡng bằng 5. Cho nên ta có thể khẳng định rằng không tồn tại vấn đề (đa) cộng tuyến trong mô hình đo lường nguyên nhân trong ví dụ của chúng ta.

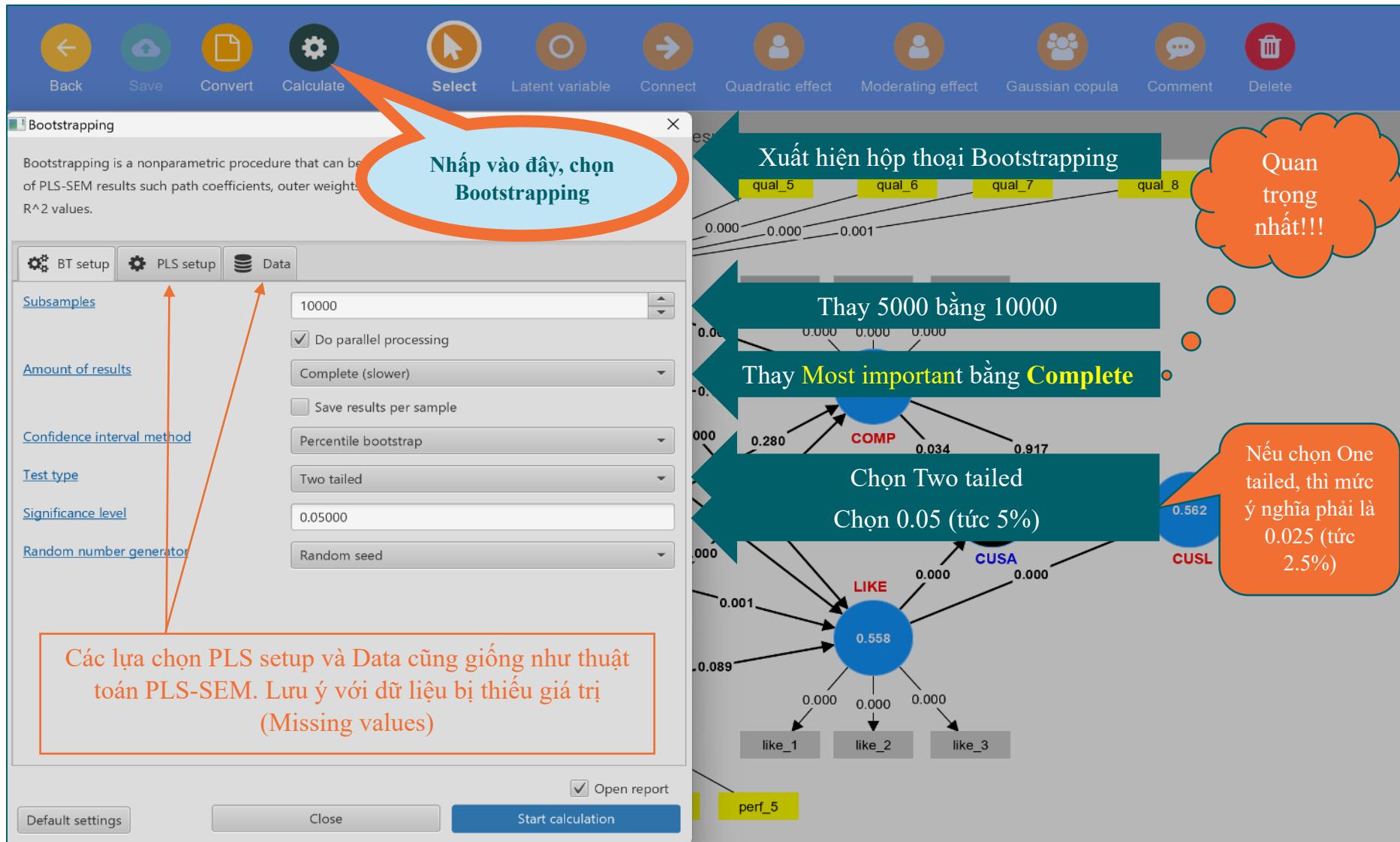
Bước 3: Đánh giá mức ý nghĩa và sự phù hợp của các biến chỉ báo nguyên nhân (Significance and relevance of the formative indicators).

Bước này là đánh giá sự đóng góp của mỗi biến chỉ báo nguyên nhân. Nhớ lại ở phương trình (2) rằng các trọng số ngoài (*w*) là kết quả hồi quy bội với điểm số biến tiềm ẩn là biến phụ thuộc và các biến chỉ báo nguyên nhân là các biến độc lập. Do khi xác định ước lượng mô hình ta chọn loại kết quả (**Type of results**) là chuẩn hóa (Standardized), nên giá trị của các ước lượng trọng số ngoài là chuẩn hóa và vì thế chúng có thể được so sánh với nhau. Trọng số nào lớn hơn thì biến chỉ báo đó đóng góp nhiều hơn vào cấu trúc; hoặc có tầm quan trọng tương đối cao hơn trong cấu trúc. Lưu ý là các giá trị ước lượng của các trọng số ngoài của các biến chỉ báo nguyên nhân trong các mô hình đo lường nguyên nhân thường nhỏ hơn các giá trị ước lượng của các hệ số tải ngoài (*t*) của các biến chỉ báo kết quả trong các mô hình đo lường kết quả (như ở các **Hình 39-41**).

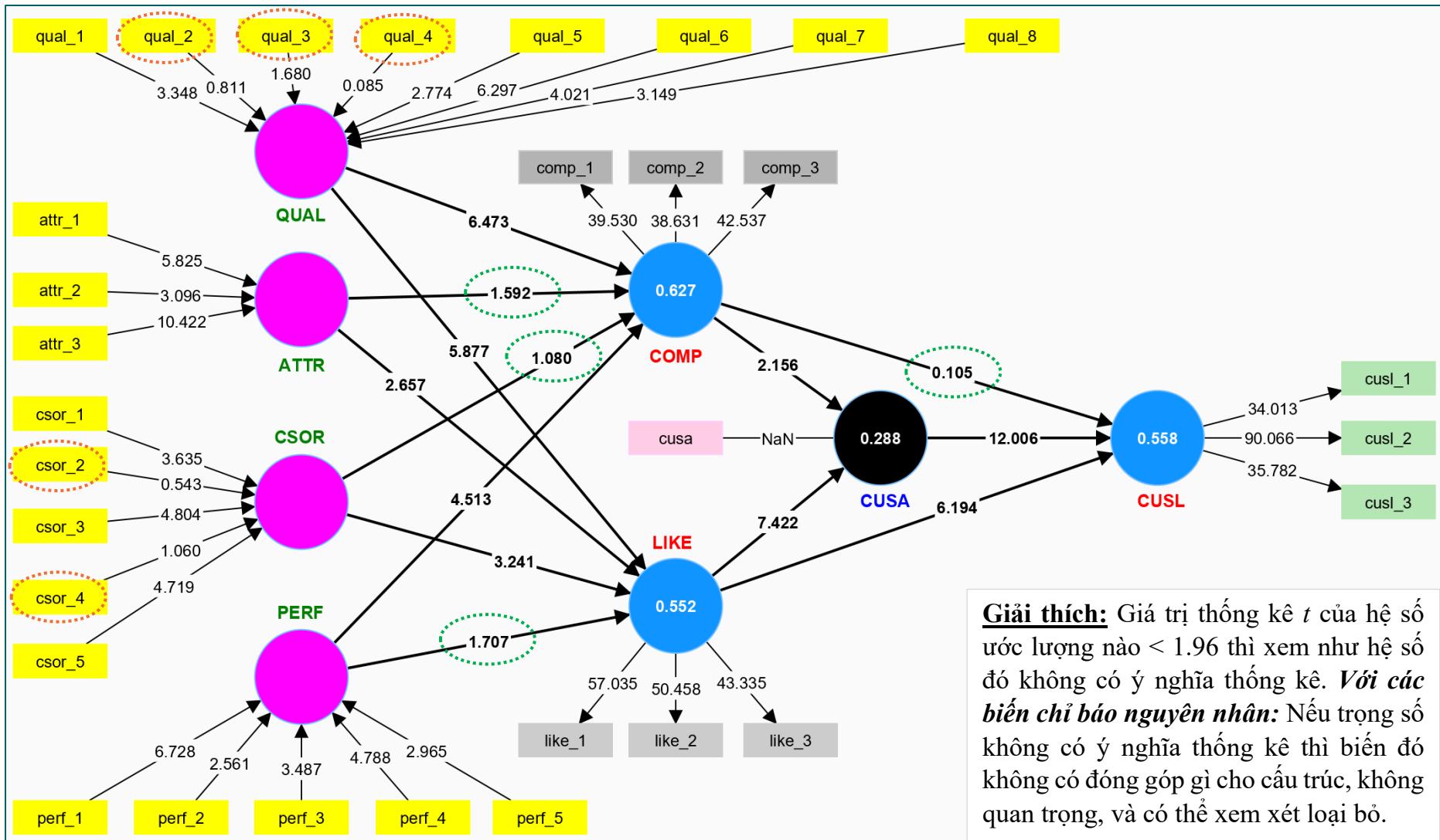
Vậy làm thế nào để biết một biến chỉ báo nguyên nhân **thực sự** đóng góp vào việc hình thành cấu trúc? Để trả lời câu hỏi này ta phải kiểm định xem các trọng số ngoài trong các mô hình đo lường nguyên nhân có khác không ($\neq 0$) một cách có ý nghĩa thống kê hay không. Như vậy, ta phải biết ít nhất một trong ba thông tin sau đây: giá trị xác suất (p-values), giá trị thống kê kiểm định (t-stats), và khoảng tin cậy (confidence interval). Như đã biết ở mô hình đơn giản, các kết quả ước lượng từ thuật toán PLS-SEM (như các **Hình 39-41**) không thể cho biết các thông tin này. Chính vì thế ta phải thực hiện lại việc ước lượng mô hình mở rộng này bằng kỹ thuật mô phỏng/bootstrapping.

Một số lưu ý cần nhắc lại khi thực hiện mô phỏng:

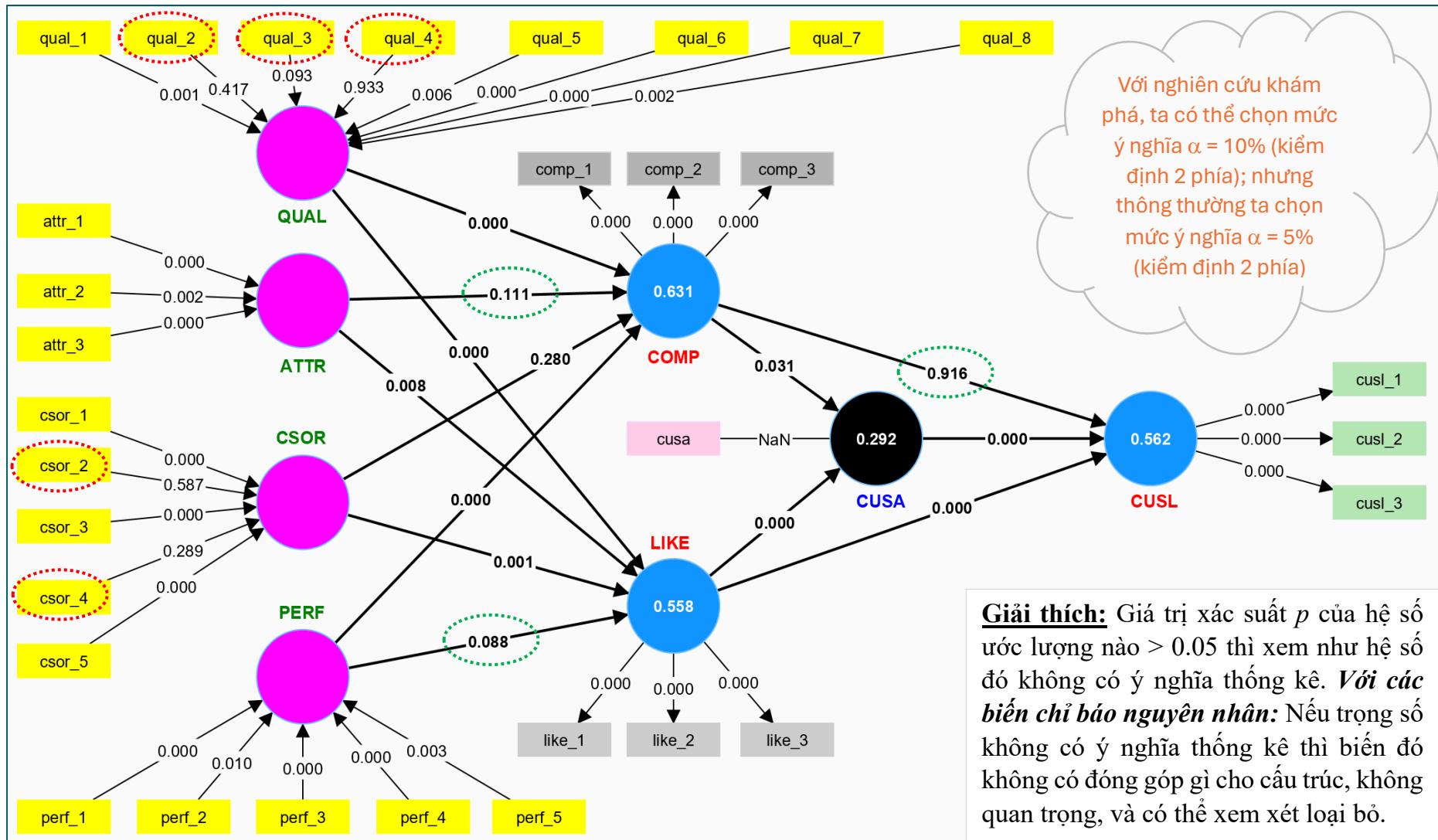
- Chọn mô hình (nhấp đúp vào **PLS Extended model** để mở sơ đồ mô hình);
- Calculate → Bootstrapping (như **Hình 44**) → Xem các ghi chú quan trọng!
- Start calculation
- Xem các kết quả ở các **Hình 45-46**.



Hình 44. Lưu ý ước lượng mô hình PLS-SEM bằng kỹ thuật mô phỏng.



Hình 45. Kết quả ước lượng mô hình mở rộng về lòng trung thành của khách hàng bằng kỹ thuật mô phỏng (Các con số trên các đường mũi tên là các giá trị **thống kê t** ; các con số trong các hình tròn là các giá trị **R^2 điều chỉnh**).



Hình 46. Kết quả ước lượng mô hình mở rộng về lòng trung thành của khách hàng bằng kỹ thuật mô phỏng (Các con số trên các đường mũi tên là các giá trị xác suất p ; các con số trong các hình tròn là các giá trị R^2).

Ngoài ra, ta cũng có thể vào **Final results → Outer weights → Means, STDEV, T values, p values** (nên sắp xếp theo giá trị giảm dần theo T statistics hoặc p values). Kết quả như **Bảng 6** cho thấy các biến có **p values** hiện màu đỏ là những biến không có ý nghĩa thống kê; đó là các biến *qual_4*, *csor_2*, *qual_2*, *csor_4*, và *qual_3*.

Bảng 6. Kết quả ước lượng mô phỏng cho các trọng số ngoài (*w*).

	Original sample (O)	Sample mean (M)	Standard deviation (STDEV)	T statistics (O/STD... ▲	P values
<i>qual_4</i> → QUAL	-0.005	-0.003	0.054	0.085	0.933
<i>csor_2</i> → CSOR	0.037	0.036	0.068	0.543	0.587
<i>qual_2</i> → QUAL	0.041	0.039	0.051	0.811	0.417
<i>csor_4</i> → CSOR	0.080	0.078	0.076	1.060	0.289
<i>qual_3</i> → QUAL	0.106	0.104	0.063	1.680	0.093
<i>perf_2</i> → PERF	0.177	0.177	0.069	2.561	0.010
<i>qual_5</i> → QUAL	0.160	0.159	0.058	2.774	0.006
<i>perf_5</i> → PERF	0.199	0.196	0.067	2.965	0.003
<i>attr_2</i> → ATTR	0.201	0.200	0.065	3.096	0.002
<i>qual_8</i> → QUAL	0.190	0.190	0.060	3.149	0.002
<i>qual_1</i> → QUAL	0.202	0.203	0.060	3.348	0.001
<i>perf_3</i> → PERF	0.194	0.192	0.056	3.487	0.000
<i>csor_1</i> → CSOR	0.306	0.304	0.084	3.635	0.000
<i>qual_7</i> → QUAL	0.229	0.227	0.057	4.021	0.000
<i>csor_5</i> → CSOR	0.416	0.414	0.088	4.719	0.000
<i>perf_4</i> → PERF	0.340	0.337	0.071	4.788	0.000
<i>csor_3</i> → CSOR	0.406	0.404	0.085	4.804	0.000
<i>attr_1</i> → ATTR	0.414	0.413	0.071	5.825	0.000
<i>qual_6</i> → QUAL	0.398	0.395	0.063	6.297	0.000
<i>perf_1</i> → PERF	0.468	0.468	0.070	6.728	0.000
<i>attr_3</i> → ATTR	0.658	0.656	0.063	10.422	0.000
<i>comp_2</i> ← COMP	0.365	0.366	0.017	21.599	0.000
<i>comp_1</i> ← COMP	0.469	0.468	0.022	21.677	0.000
<i>cusl_1</i> ← CUSL	0.369	0.369	0.016	22.479	0.000
<i>cusl_3</i> ← CUSL	0.365	0.365	0.016	23.515	0.000

Kết quả này cho thấy khi có nhiều biến chỉ số nguyên nhân được dùng để đo lường một cấu trúc (như **QUAL** chẳng hạn), thì có khả năng cao hơn là một hay nhiều biến chỉ báo sẽ có trọng số ngoài thấp hoặc thậm chí là không có ý nghĩa thống kê. Theo [Hair và cộng sự \(2022: trang 149\)](#), khi số biến chỉ số nguyên nhân được giả định là không tương quan với nhau, thì trọng số ngoài trung bình là $1/\sqrt{n}$, với n là số biến chỉ báo nguyên nhân. Không chỉ trọng số ngoài cao nhất có thể giảm khi số biến chỉ báo nguyên nhân tăng, mà giá trị trung bình của các trọng số ngoài cũng giảm một cách có ý nghĩa thống kê khi số lượng biến chỉ báo nguyên nhân tăng lên. Vì thế, số biến chỉ báo nguyên nhân sẽ trở nên không có ý nghĩa thống kê tăng thêm khi có nhiều biến hơn được dùng để đo lường một cấu trúc. Lưu ý rằng, các biến chỉ báo nguyên nhân không nên bị loại bỏ chỉ dựa trên kết quả thống kê. Trước khi loại bỏ một biến chỉ báo nguyên nhân, nhà nghiên cứu nên đánh giá sự phù hợp của nó về mặt giá trị nội dung.

Cuối cùng, hãy trình bày tóm tắt các kết quả đánh giá mô hình đo lường nguyên nhân như đề xuất ở **Bảng 7**.

Bảng 7. Tóm tắt kết quả chính của mô hình đo lường nguyên nhân.

Formative Constructs	Formative Indicators	Outer Weights (Outer Loadings)	t Value	p Value	95% Percentile Confidence Interval	Significance ^a ($p < 0.05$)?
ATTR	attr_1	0.414 (0.755)	5.851	0.000	[0.270, 0.547]	Yes
	attr_2	0.201 (0.506)	3.088	0.002	[0.073, 0.329]	Yes
	attr_3	0.658 (0.891)	10.522	0.000	[0.529, 0.775]	Yes
CSOR	csor_1	0.306 (0.771)	3.640	0.000	[0.140, 0.467]	Yes
	csor_2	0.037 (0.571)	0.554	0.580	[-0.095, 0.170]	No
	csor_3	0.406 (0.838)	4.729	0.000	[0.229, 0.564]	Yes
	csor_4	0.080 (0.617)	1.057	0.290	[-0.071, 0.2223]	No
	csor_5	0.416 (0.848)	4.706	0.000	[0.243, 0.588]	Yes
PERF	perf_1	0.468 (0.846)	6.724	0.000	[0.326, 0.599]	Yes
	perf_2	0.177 (0.690)	2.588	0.010	[0.040, 0.309]	Yes
	perf_3	0.194 (0.573)	3.439	0.001	[0.080, 0.300]	Yes
	perf_4	0.340 (0.717)	4.797	0.000	[0.199, 0.477]	Yes
	perf_5	0.199 (0.638)	2.963	0.003	[0.071, 0.334]	Yes
QUAL	qual_1	0.202 (0.741)	3.323	0.001	[0.083, 0.323]	Yes
	qual_2	0.041 (0.570)	0.796	0.426	[-0.061, 0.142]	No
	qual_3	0.106 (0.749)	1.686	0.092	[-0.019, 0.227]	No
	qual_4	-0.005 (0.664)	0.083	0.934	[-0.111, 0.105]	No
	qual_5	0.160 (0.787)	2.785	0.005	[0.042, 0.269]	Yes
	qual_6	0.398 (0.856)	6.162	0.000	[0.269, 0.525]	Yes
	qual_7	0.229 (0.722)	3.997	0.000	[0.116, 0.340]	Yes
	qual_8	0.190 (0.627)	3.160	0.002	[0.073, 0.308]	Yes

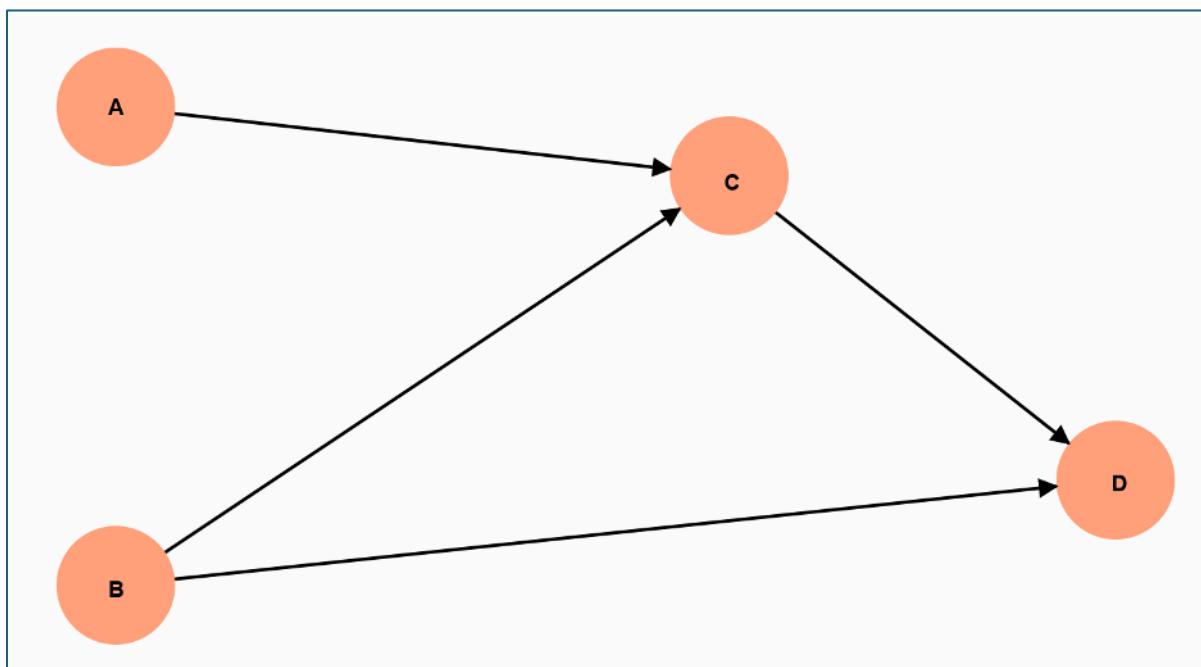
Nguồn: Hair và cộng sự (2022: trang 181).

6.4 Đánh giá mô hình cấu trúc

Một khi ta đã xác nhận các thước đo cấu trúc là đáng tin cậy và chuẩn xác, bước tiếp theo là đánh giá các kết quả của mô hình cấu trúc. Mục đích của việc đánh giá các kết quả của mô hình cấu trúc là giúp ta xác định khả năng của mô hình trong việc giải thích và dự đoán một hay nhiều cấu trúc/biến tiềm ẩn quan tâm có liên quan đến các mục tiêu nghiên cứu. Các nội dung chính của việc đánh giá này thường bao gồm: 1) Kiểm định (đa) cộng tuyến; 2) Đánh giá sự ý nghĩa và phù hợp của các mối quan hệ trong mô hình cấu trúc, tức là đánh giá các hệ số đường dẫn; 3) Đánh giá năng lực giải thích và dự đoán của mô hình cấu trúc.

Bước 1: Đánh giá sự cộng tuyến của mô hình cấu trúc (Collinearity of the structural model).

Để đánh giá (đa) cộng tuyến, ta áp dụng các thước đo hoàn toàn giống với các thước đo đã áp dụng cho các mô hình đo lường nguyên nhân (đó là TOL và VIF). Do trong mô hình cấu trúc, nhất là mô hình phức tạp với nhiều biến tiềm ẩn nội sinh, nên ta cần phân tích mỗi nhóm cấu trúc dự đoán (**predictor constructs**) một cách riêng biệt cho từng nhóm cấu trúc của mô hình cấu trúc. Ví dụ, **Hình 47** cho thấy các cấu trúc A và B cùng giải thích cho cấu trúc C. Tương tự, các cấu trúc B và C cùng giải thích cho cấu trúc D. Như vậy, ta cần kiểm tra liệu có mức độ cộng tuyến cao giữa mỗi nhóm biến dự đoán; đó là, A và B cũng như B và C.



Hình 47. Khả năng cộng tuyến trong mô hình cấu trúc.

Tương tự như đánh giá các mô hình đo lường nguyên nhân, các giá trị VIF trong các cấu trúc dự đoán phải dưới 5 và tốt nhất là dưới 3 để đảm bảo rằng (đa) cộng tuyến

không có ảnh hưởng đáng kể đến các giá trị ước lượng của mô hình cấu trúc. Nếu VIF chỉ ra mức độ quan trọng của (đa) cộng tuyến, thì ta nên xem xét loại bỏ các cấu trúc, hợp nhất các cấu trúc dự đoán thành một cấu trúc duy nhất hoặc tạo các cấu trúc bậc cao hơn để xử lý các vấn đề (đa) cộng tuyến.

Trở lại với mô hình mở rộng về lòng trung thành của khách hàng. Để kiểm định (đa) cộng tuyến trên SmartPLS 4 sau khi đã thực hiện ước lượng mô hình bằng thuật toán PLS-SEM ta tiến hành như sau:

- Quality Criteria → Collinearity Statistic (VIF)
 - Inner model – Matrix

PLS-SEM algorithm		Collinearity statistics(VIF)- Inner model - Matrix							
		ATTR	COMP	CSOR	CUSA	CUSL	LIKE	PERF	QUAL
<input checked="" type="checkbox"/>	Graphical		2.122				2.122		
<input checked="" type="checkbox"/>	Final results				1.686	1.716			
<input checked="" type="checkbox"/>	Quality criteria		2.083				2.083		
<input checked="" type="checkbox"/>	R-square					1.412			
<input checked="" type="checkbox"/>	f-square								
<input checked="" type="checkbox"/>	Construct reliability and validity								
<input checked="" type="checkbox"/>	Discriminant validity								
<input checked="" type="checkbox"/>	Collinearity statistics (VIF)				1.686	1.954			
<input checked="" type="checkbox"/>	Outer model - List								
<input checked="" type="checkbox"/>	Inner model - Matrix		2.889				2.889		
<input checked="" type="checkbox"/>	Inner model - List		3.487				3.487		

Hình 48. Các giá trị VIF trong mô hình cấu trúc của mô hình mở rộng.

Trong **Hình 48**, các cấu trúc nội sinh được thể hiện theo các cột và các cấu trúc ngoại sinh (tức là các biến dự đoán) được thể hiện theo các dòng. Ta sẽ đánh giá (đa) cộng tuyến cho các nhóm cấu trúc dự đoán sau đây: **1)** ATTR, CSOR, PERF, và QUAL là các biến dự đoán cho COMP và LIKE; **2)** COMP và LIKE là các biến dự đoán cho CUSA; và **3)** COMP, LIKE, và CUSA là các biến dự đoán cho CUSL. Kết quả cho thấy tất cả các giá trị VIF đều dưới giá trị ngưỡng bằng 5 và với một ngoại lệ (đó là biến dự đoán QUAL) là lớn hơn 3. Vì giá trị VIF của QUAL gần bằng 3, nên ta kết luận rằng (đa) cộng tuyến giữa các cấu trúc dự đoán là không phải vấn đề quan trọng trong mô hình cấu trúc, và ta có thể tiếp tục phân tích các kết quả ước lượng của mô hình.

Bước 2: Đánh giá mức ý nghĩa và sự phù hợp của các mối quan hệ trong mô hình cấu trúc (Significance and relevance of the structural model relationships).

Với lựa chọn loại kết quả là “**standardized**” thì các kết quả ước lượng từ thuật toán PLS-SEM cho ta các mối quan hệ của mô hình cấu trúc (tức các hệ số đường dẫn, p) là các mối quan hệ “chuẩn hóa” giữa các cấu trúc. Các hệ số đường dẫn có các giá trị chuẩn hóa dao động trong khoảng từ -1 đến +1. Các giá trị ước lượng của hệ số đường dẫn gần +1 thể hiện các mối quan hệ dương mạnh (và ngược lại), và thường là có ý

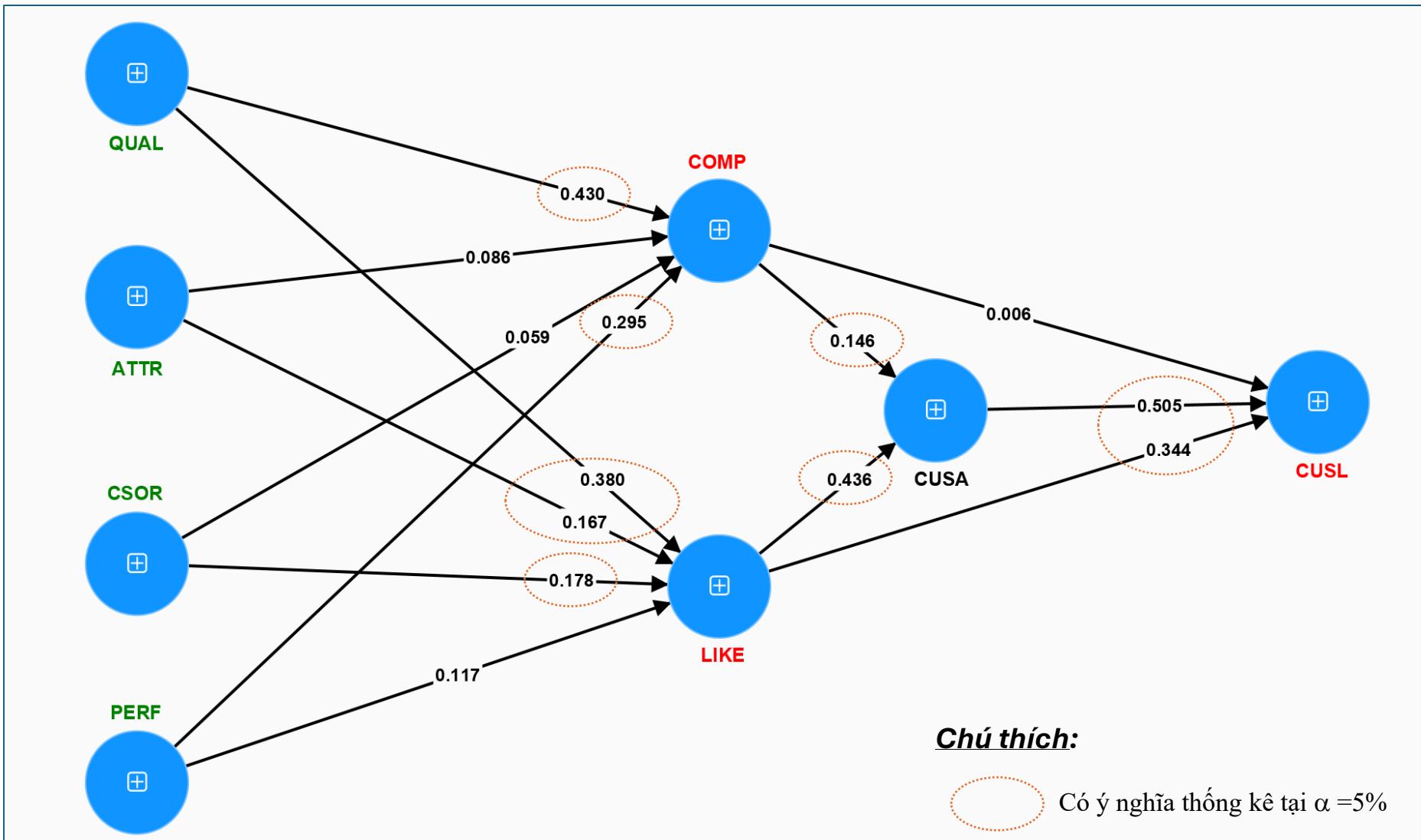
nghĩa về mặt thống kê (nghĩa là chắc chắn hệ số đường dẫn đó là khác 0). Các giá trị ước lượng của hệ số đường dẫn càng gần 0 thì các mối quan hệ càng yếu đi, và tầm quan trọng tương đối của cấu trúc (dự đoán) trong việc giải thích một cấu trúc khác trong mô hình cấu trúc. Các giá trị ước lượng của hệ số đường dẫn càng gần 0 thường là không có ý nghĩa thống kê.

Để kiểm định mức ý nghĩa của các hệ số đường dẫn thì ta phải có các thông tin về sai số chuẩn, giá trị thống kê kiểm định, giá trị xác suất, và khoảng tin cậy của các hệ số đường dẫn. Như đã giải thích ở trên, ta phải thực hiện ước lượng lại mô hình bằng kỹ thuật mô phỏng (**bootstrapping**). Các sai số chuẩn từ kết quả mô phỏng (sai số chuẩn bootstrap) giúp ta tính các giá trị thống kê t và xác suất p cho tất cả các hệ số đường dẫn. Khi một giá trị thống kê t tính toán (**computed t value**) lớn hơn giá trị thống kê t phê phán (**critical t value**), thì ta kết luận là hệ số ước lượng có ý nghĩa thống kê ở một mức ý nghĩa nhất định (α). Thông thường, các giá trị t phê phán cho một kiểm định hai phía (**two-tailed test**) là **1.65** (với $\alpha = 10\%$), **1.96** (với $\alpha = 5\%$), và **2.57** (với $\alpha = 1\%$). Các giá trị t phê phán cho một kiểm định một phía (**one-tailed test**) là **1.28** (với $\alpha = 10\%$), **1.65** (với $\alpha = 5\%$), và **2.33** (với $\alpha = 1\%$). Khi thực hiện ước lượng bằng **Bootstrapping** trên SmartPLS 4, ta cần phải xác định loại kiểm định một hay hai phía để biết các kết quả hệ số ước lượng thích hợp với loại kiểm định nào.

Tuy nhiên, hầu hết các nhà nghiên cứu sử dụng các giá trị xác suất p (**p values**) để đánh giá mức ý nghĩa của các hệ số ước lượng. Một giá trị xác suất p nhỏ hơn một mức ý nghĩa α được chọn thì hệ số ước lượng đó được cho là có ý nghĩa thống kê ở mức ý nghĩa α . Tương tự như giá trị thống kê t , ta cũng cần biết loại kiểm định được chọn khi thực hiện ước lượng bằng Bootstrapping là một hay hai phía để có kết luận đúng đắn. Nhắc lại rằng, phổ biến nhất là ta nên chọn loại kiểm định hai phía.

Một cách thứ ba ít phổ biến hơn là sử dụng khoảng tin cậy để xác định một hệ số ước lượng có ý nghĩa thống kê hay không. Theo cách này thì nếu một khoảng tin cậy cho một hệ số ước lượng không chứa giá trị 0, thì ta bác bỏ giả thuyết không [tức là ($H_0: p_k = 0$ bị bác bỏ)]; và điều này hàm ý rằng hệ số ước lượng đó có ý nghĩa thống kê (tức khác 0 một cách có ý nghĩa thống kê).

Theo Hair và cộng sự (2022: trang 193), do các hệ số đường dẫn ở dạng chuẩn hóa (với lựa chọn “standardized” khi thực hiện ước lượng), nên chúng có thể được giải thích dạng so sánh tầm quan trọng tương đối của cấu trúc dự đoán này với một cấu trúc dự đoán khác. Nếu một hệ số đường dẫn lớn hơn hệ số đường dẫn khác, thì ảnh hưởng của nó lên biến tiềm ẩn nội sinh là lớn hơn. Cách giải thích các hệ số đường dẫn với các hệ số ước lượng chuẩn hóa trong mô hình cấu trúc như sau: *Một thay đổi của cấu trúc ngoại sinh bằng một độ lệch chuẩn (standard deviation) làm thay đổi độ lệch chuẩn của cấu trúc ngoại sinh bằng độ lớn của hệ số đường dẫn khi mọi thứ khác (tức tất cả cấu trúc khác trong mô hình cấu trúc) được giữ nguyên không đổi.* Nếu hệ số đường dẫn có ý nghĩa thống kê thì giá trị ước lượng của nó cho biết mức độ mà cấu trúc ngoại sinh đó quan hệ với cấu trúc nội sinh.



Hình 48. Hệ số đường dẫn của mô hình mở rộng về lòng trung thành của khách hàng.

Như đã được trình bày ở **Mục 3**, ta thường quan tâm không chỉ là ảnh hưởng trực tiếp (**direct effect**) của một cấu trúc ngoại sinh lên một cấu trúc nội sinh mà còn là các ảnh gián tiếp (**indirect effects**) lên cấu trúc nội sinh đó thông qua một hay nhiều cấu trúc trung gian (**mediating constructs**). Tổng của các ảnh hưởng trực tiếp và gián tiếp được gọi là tổng ảnh hưởng (**total effect**). Kết quả ước lượng bằng kỹ thuật Bootstraping cũng cho ta các kiểm định ý nghĩa của các loại ảnh hưởng này. Chúng ta sẽ giải thích kỹ hơn về các ảnh hưởng trung gian ở phần sau.

Trở lại với ví dụ mô hình mở rộng về lòng trung thành của khách hàng với các hệ số đường dẫn được trình bày ở các **Hình 39-41**, và mức ý nghĩa ở các **Hình 45-46**. Sử dụng giá trị xác suất p như ở **Hình 46** ta nhận thấy rằng các hệ số đường dẫn sau đây không có ý nghĩa thống kê ở mức ý nghĩa $\alpha = 5\%$: pATTR > COMP, pCSR > COMP, pPERF > LIKE, và pCOMP > CUSL. Điều này có nghĩa là các hệ số đường dẫn này là bằng 0. Ngoại trừ pCOMP > CUSA, thì tất cả các hệ số đường dẫn còn lại đều có ý nghĩa thống kê rất cao (ở mức ý nghĩa $\alpha = 1\%$). Về mức độ quan trọng của các cấu trúc thì ta có thể sử dụng một trong các **Hình 39-41**. Tuy nhiên, để việc phân tích trở nên dễ dàng hơn thì ta trở lại kết quả ước lượng trên SmartPLS thực hiện một số chỉnh sửa như ẩn các biến chỉ báo, chuyển vị trí các hệ số đường dẫn có ý nghĩa thống kê gần sát với các cấu trúc nội sinh như **Hình 48**. Kết quả cho ta thấy như sau:

- Với biến nội sinh COMP: $pQUAL > COMP = 0.430 > pPERF > COMP = 0.295 \Rightarrow$ Cảm nhận về chất lượng sản phẩm và dịch vụ của công ty tác động nhiều hơn đến cảm nhận về năng lực của công ty so với cảm nhận về hiệu quả kinh tế và quản trị của công ty.
- Với biến nội sinh LIKE: $pQUAL > LIKE = 0.380 > pCSR > LIKE = 0.178 \sim pATTR > LIKE = 0.176 \Rightarrow$ Cảm nhận về chất lượng sản phẩm và dịch vụ của công ty tác động nhiều hơn đến cảm nhận về sự yêu thích so với cảm nhận về trách nhiệm xã hội của công ty và cảm nhận về sự thu hút của công ty.
- Với biến nội sinh CUSA: $pLIKE > CUSA = 0.436 > pCOMP > CUSA = 0.146 \Rightarrow$ Cảm nhận về sự yêu thích của công ty ảnh hưởng nhiều hơn đến cảm nhận về sự hài lòng của khách hàng so với cảm nhận về năng lực của công ty.
- Với biến nội sinh CUSL: $pCUSA > CUSL = 0.505 > pLIKE > CUSL = 0.344 \Rightarrow$ Cả hai biến về sự hài lòng của khách hàng và cảm nhận về sự yêu thích của công ty đều có ảnh hưởng cao đến lòng trung thành của khách hàng; tuy nhiên, tầm quan trọng của sự hài lòng của khách hàng thì cao hơn.
- Cảm nhận về sự yêu thích là động lực chính của cả sự hài lòng và lòng trung thành của khách hàng chứ không phải cảm nhận về năng lực của công ty.

Bây giờ ta phân tích tổng ảnh hưởng từ mô hình cấu trúc. Cụ thể, ta sẽ đánh giá mỗi cấu trúc nguyên nhân (QUAL, ATTR, CSOR, và PERF) ảnh hưởng mạnh như thế nào đến biến mục tiêu của nghiên cứu (tức là biến CUSL) thông qua các cấu trúc trung gian COMP, LIKE, và CUSA. Để thực hiện đánh giá này, từ kết quả ước lượng mô hình mở rộng bằng thuật toán PLS-SEM, ta chọn: **Final reslts → Total effects**. Kết quả được trình bày trong **Hình 49**.

PLS-SEM algorithm		Total effects - Matrix							
		ATTR	COMP	CSOR	CUSA	CUSL	LIKE	PERF	QUAL
▼ Graphical	Graphical output		0.086		0.085	0.101	0.167		
▼ Final results		ATTR			0.146	0.079			
▶ Path coefficients		COMP							
▶ Indirect effects		CSOR	0.059		0.086	0.105	0.178		
▼ Total effects		CUSA					0.505		
○ Matrix		CUSL							
○ List		LIKE			0.436	0.564			
▶ Outer loadings		PERF	0.295		0.094	0.089	0.117		
		QUAL	0.430		0.228	0.248	0.380		

Hình 49. Tổng ảnh hưởng của các biến cấu trúc dự đoán lên cấu trúc nội sinh.

Trong ma trận ở **Hình 49**, mỗi cột thể hiện một cấu trúc mục tiêu (**target construct**, thường là các cấu trúc nội sinh); và các dòng thể hiện các cấu trúc tiền vị (**predecessor constructs**, nghĩa là các cấu trúc ảnh hưởng lên cấu trúc mục tiêu, có thể chỉ đóng vai trò cấu trúc ngoại sinh hoặc vừa ngoại sinh vừa nội sinh). Ví dụ, với cấu trúc mục tiêu là lòng trung thành (CUSL), ta thấy rằng trong bốn cấu trúc nguyên nhân (hay cấu trúc ngoại sinh) thì cảm nhận về chất lượng sản phẩm và dịch vụ (QUAL) có tổng ảnh hưởng mạnh nhất đến lòng trung thành (0.248), tiếp theo là trách nhiệm xã hội của công ty (CSOR, 0.105), sự thu hút (ATTR, 0.101), và sau cùng là hiệu quả kinh tế và quản trị (PERF, 0.089). Con số **0.248** được tính như thế nào? Nhìn vào **Hình 48** ta thấy giá trị tổng ảnh hưởng này chỉ bao gồm tổng ảnh hưởng gián tiếp (**total indirect effects**) bởi vì QUAL không có đường dẫn trực tiếp đến CUSL. Tổng ảnh hưởng gián tiếp này thông qua các cấu trúc trung gian: COMP, LIKE, và CUSA.

$$\begin{aligned}
 0.248 &= 0.430 \times 0.006 && \Rightarrow \text{ảnh hưởng gián tiếp qua COMP} \\
 &+ 0.430 \times 0.146 \times 0.505 && \Rightarrow \text{ảnh hưởng gián tiếp qua COMP và CUSA} \\
 &+ 0.380 \times 0.436 \times 0.505 && \Rightarrow \text{ảnh hưởng gián tiếp qua LIKE và CUSA} \\
 &+ 0.380 \times 0.344 && \Rightarrow \text{ảnh hưởng gián tiếp qua LIKE}
 \end{aligned}$$

Cấu trúc LIKE vừa có ảnh hưởng trực tiếp vừa có ảnh hưởng gián tiếp lên CUSL thông qua CUSA. Giá trị tổng ảnh hưởng của LIKE lên CUSL là **0.564**. Con số này được tính như sau:

$$\begin{aligned}
 0.564 &= 0.344 && \Rightarrow \text{ảnh hưởng trực tiếp (hệ số đường dẫn)} \\
 &+ 0.436 \times 0.505 && \Rightarrow \text{ảnh hưởng gián tiếp qua CUSA}
 \end{aligned}$$

Tuy nhiên, để biết các (tổng) ảnh hưởng gián tiếp và tổng ảnh hưởng có ý nghĩa thống kê hay không; tức giá trị ảnh hưởng có khác 0 một cách có ý nghĩa thống kê, ta cần phải dựa vào kết quả mô phỏng với kỹ thuật Bootstrapping. Kết quả như các **Bảng 8 và 9**.

Bảng 8. Tổng ảnh hưởng gián tiếp của các cấu trúc dự đoán.

	Original sample ... ▾	Sample mean (M)	Standard deviation (STDEV)	T statistics (t /STDEVI)	P values
QUAL -> CUSL	0.248	0.253	0.043	5.821	0.000
QUAL -> CUSA	0.228	0.231	0.037	6.175	0.000
LIKE -> CUSL	0.220	0.220	0.035	6.347	0.000
CSOR -> CUSL	0.105	0.106	0.034	3.092	0.001
ATTR -> CUSL	0.101	0.099	0.037	2.721	0.003
PERF -> CUSA	0.094	0.096	0.039	2.427	0.008
PERF -> CUSL	0.089	0.091	0.044	2.009	0.022
CSOR -> CUSA	0.086	0.087	0.028	3.127	0.001
ATTR -> CUSA	0.085	0.083	0.030	2.811	0.002
COMP -> CUSL	0.074	0.074	0.035	2.084	0.019

Bảng 9. Tổng ảnh hưởng của các cấu trúc dự đoán.

	Original sample ... ▾	Sample mean (M)	Standard deviation (STDEV)	T statistics (t /STDEVI)	P values
LIKE -> CUSL	0.564	0.565	0.059	9.480	0.000
CUSA -> CUSL	0.505	0.506	0.042	12.027	0.000
LIKE -> CUSA	0.436	0.436	0.059	7.347	0.000
QUAL -> COMP	0.430	0.430	0.066	6.483	0.000
QUAL -> LIKE	0.380	0.386	0.065	5.851	0.000
PERF -> COMP	0.295	0.300	0.066	4.468	0.000
QUAL -> CUSL	0.248	0.253	0.043	5.821	0.000
QUAL -> CUSA	0.228	0.231	0.037	6.175	0.000
CSOR -> LIKE	0.178	0.179	0.057	3.135	0.001
ATTR -> LIKE	0.167	0.163	0.063	2.641	0.004
COMP -> CUSA	0.146	0.146	0.068	2.149	0.016
PERF -> LIKE	0.117	0.119	0.070	1.670	0.047
CSOR -> CUSL	0.105	0.106	0.034	3.092	0.001
ATTR -> CUSL	0.101	0.099	0.037	2.721	0.003
PERF -> CUSA	0.094	0.096	0.039	2.427	0.008
PERF -> CUSL	0.089	0.091	0.044	2.009	0.022
ATTR -> COMP	0.086	0.085	0.054	1.587	0.056
CSOR -> CUSA	0.086	0.087	0.028	3.127	0.001
ATTR -> CUSA	0.085	0.083	0.030	2.811	0.002
COMP -> CUSL	0.079	0.079	0.067	1.186	0.118
CSOR -> COMP	0.059	0.062	0.054	1.090	0.138

Các **Bảng 8** và **9** được thực hiện như sau (từ kết quả mô phỏng bằng Bootstrapping): **Final results → Total indirect effects (hoặc Total effects) → Mean, STDEV, T values, p values**. Các giá trị xác suất *p* hiển thị màu đỏ cho biết những tổng ảnh hưởng (gián tiếp) nào không có ý nghĩa thống kê ở mức ý nghĩa $\alpha = 5\%$. Tất cả các tổng ảnh hưởng gián tiếp (**Bảng 8**) đều có ý nghĩa thống kê. **Bảng 9** cho thấy có ba tổng ảnh hưởng không có ý nghĩa thống kê; đó là, ATTR → COMP, COMP → CUSL, và CSOR → COMP. Từ kết quả mô phỏng, ta có thể sort giá trị theo hướng giảm dần hoặc tăng dần để có cái nhìn rõ ràng hơn về tầm quan trọng của mỗi cấu trúc dự đoán trong mô hình.

Khi trình bày kết quả ước lượng và kiểm định ý nghĩa của mô hình cấu trúc, ta nên lập các bản tóm tắt như gợi ý ở **Bảng 10 và 11**.

Bảng 10. Các kết quả kiểm định ý nghĩa các hệ số đường dẫn.

	Path Coefficients	t Values	p Values	95% Confidence Intervals	Significance ^a ($p < 0.05$)?
ATTR → COMP	0.086	1.579	0.114	[-0.019, 0.194]	No
ATTR → LIKE	0.167	2.666	0.008	[0.037, 0.284]	Yes
COMP → CUSA	0.146	2.173	0.030	[0.015, 0.275]	Yes
COMP → CUSL	0.006	0.103	0.918	[-0.101, 0.113]	No
CSOR → COMP	0.059	1.086	0.278	[-0.044, 0.169]	No
CSOR → LIKE	0.178	3.200	0.001	[0.067, 0.288]	Yes
CUSA → CUSL	0.505	11.960	0.000	[0.419, 0.583]	Yes
LIKE → CUSA	0.436	7.381	0.000	[0.316, 0.550]	Yes
LIKE → CUSL	0.344	6.207	0.000	[0.236, 0.454]	Yes
PERF → COMP	0.295	4.457	0.000	[0.168, 0.428]	Yes
PERF → LIKE	0.117	1.684	0.092	[-0.013, 0.259]	No
QUAL → COMP	0.430	6.400	0.000	[0.300, 0.562]	Yes
QUAL → LIKE	0.380	5.834	0.000	[0.260, 0.515]	Yes

Nguồn: Hair và cộng sự (2022: trang 217).

Bảng 11. Các kết quả kiểm định của tổng ảnh hưởng.

	Total Effect	t Values	p Values	95% Confidence Intervals	Significance ^a ($p < 0.05$)?
ATTR → CUSA	0.085	2.817	0.005	[0.024, 0.144]	Yes
ATTR → CUSL	0.101	2.740	0.006	[0.026, 0.172]	Yes
CSOR → CUSA	0.086	3.142	0.002	[0.034, 0.143]	Yes
CSOR → CUSL	0.105	3.107	0.002	[0.042, 0.175]	Yes
PERF → CUSA	0.094	2.454	0.014	[0.021, 0.172]	Yes
PERF → CUSL	0.089	2.034	0.042	[0.005, 0.177]	Yes
QUAL → CUSA	0.228	6.189	0.000	[0.161, 0.306]	Yes
QUAL → CUSL	0.248	5.788	0.000	[0.171, 0.339]	Yes

Nguồn: Hair và cộng sự (2022: trang 219).

Bước 3: Đánh giá sức mạnh giải thích của mô hình (Model's explanatory power).

Theo Hair và cộng sự (2022: trang 195), sức mạnh giải thích của mô hình liên quan đến khả năng phù hợp với dữ liệu có sẵn của nó. Thước đo được sử dụng phổ biến nhất về sức mạnh giải thích của mô hình cấu trúc là hệ số xác định (R^2); được tính như bình phương của hệ số tương quan giữa các giá trị thực và giá trị ước lượng của một cấu trúc nội sinh cụ thể. Hệ số xác định này thể hiện các ảnh hưởng kết hợp của các biến tiềm ẩn ngoại sinh lên một biến tiềm ẩn nội sinh cụ thể. Nói cách khác, hệ số xác định thể hiện lượng phương sai trong một cấu trúc nội sinh được giải thích bởi tất cả các cấu trúc ngoại sinh liên quan đến nó. Giá trị R^2 nằm trong khoảng từ 0 đến 1; trong đó, các mức càng cao chứng tỏ sức mạnh giải thích càng lớn. Lưu ý, R^2 là một hàm của số các cấu trúc dự đoán, nên số cấu trúc dự đoán càng nhiều thì giá trị R^2 càng cao. Vì thế, R^2 luôn luôn được giải thích theo bối cảnh của nghiên cứu, dựa trên các giá trị R^2 từ các nghiên cứu liên quan và các mô hình có độ phức tạp tương tự. Chính vì vậy, khi với nhiều cấu trúc dự đoán, ta nên sử dụng các giá trị R^2 điều chỉnh để so sánh giữa các mô hình cấu trúc khác nhau.

Ngoài ra, giá trị R^2 cũng có thể được sử dụng để lượng hóa độ mạnh của các mối quan hệ trong mô hình cấu trúc bằng cách tính độ lớn ảnh hưởng f^2 (effect size). Độ lớn ảnh hưởng f^2 thể hiện sự thay đổi trong giá trị R^2 khi một cấu trúc tiền vị bị bỏ ra khỏi mô hình. Nếu việc loại bỏ một cấu trúc dẫn đến một sự sụt giảm đáng kể trong giá trị R^2 của cấu trúc nội sinh thì cấu trúc tiền vị đó được cho là có tác động đáng kể lên cấu trúc nội sinh. Độ lớn f^2 đã được định nghĩa ở [công thức \(12\)](#), [Mục 5.2](#).

Trong hình tròn ứng với các cấu trúc nội sinh (như ở [Hình 39](#)) hiển thị mặc định các giá trị R^2 từ các kết quả ước lượng với cả thuật toán PLS-SEM và Bootstrapping. Ngoài ra, ta có thể xem các giá trị R^2 sau kết quả ước lượng với thuật toán PLS-SEM bằng cách: [Quality criteria → R-square → Overview](#). Theo quy tắc chung thì ngưỡng tiêu chuẩn đối với R^2 là **0.5**. Vì vậy, các giá trị R^2 của COMP (0.631), CUSL (0.562), và LIKE (0.558) có thể được xem là vừa phải, trong khi đó giá trị R^2 của CUSA (0.292) thì khá yếu. Tương tự, ta có thể xem độ lớn ảnh hưởng f^2 (từ kết quả ước lượng với thuật toán PLS-SEM) bằng cách: [Quality criteria → f-square → Matrix](#). Kết quả như sau:

Bảng 12. Các giá trị độ lớn ảnh hưởng f^2 (màu đỏ là các ảnh hưởng không đáng kể).

	ATTR	COMP	CSOR	CUSA	CUSL	LIKE	PERF	QUAL
ATTR		0.009				0.030		
COMP				0.018	0.000			
CSOR		0.005				0.035		
CUSA					0.412			
CUSL								
LIKE				0.159	0.138			
PERF		0.082					0.011	
QUAL		0.143				0.094		

Bước 4: Đánh giá sức mạnh dự đoán của mô hình (Model's predictive power).

Sức mạnh giải thích đôi khi còn được hiểu là sức mạnh dự đoán trong mẫu (im-sample predictive power) (Chin, 2010; Rigdon, 2012; Sarstedt và cộng sự, 2014). Tuy nhiên, để một mô hình đường dẫn PLS hữu ích cho việc ra quyết định, thì mô hình cần tạo ra các phát hiện có thể khái quát hóa (Hair và Sarstedt, 2021). Tạo ra các phát hiện có thể khái quát hóa yêu cầu đánh giá liệu các kết quả có thể áp dụng không chỉ cho các dữ liệu đã được sử dụng trong quá trình ước lượng mô hình mà còn các tệp dữ liệu khác không được đưa vào quá trình ước lượng. Theo Chin và cộng sự (2020) và Sarstedt và cộng sự (2017) thì cần đánh giá sức mạnh dự đoán ngoài mẫu của mô hình (**out-of-sample predictive power**) hay đơn giản chỉ là sức mạnh dự đoán.

Thuật toán PLS_{predict} của Shmueli và cộng sự (2016) là phương pháp chủ đạo để đánh giá sức mạnh dự đoán của một mô hình đường dẫn PLS. Sự phân tích được xây dựng trên khái niệm chia tách tệp dữ liệu tổng thể thành các mẫu huấn luyện và giữ lại/thử nghiệm (**training and holdout/test samples**). Mẫu huấn luyện là một phần của tệp dữ liệu được dùng để ước lượng các tham số của mô hình (như các hệ số đường dẫn, các trọng số ngoài, và các hệ số tái ngoài). Phần còn lại của tệp dữ liệu không được sử dụng vào việc ước lượng mô hình được gọi là mẫu thử nghiệm (Danks và cộng sự, 2017). Để dự đoán các giá trị của các biến chỉ báo của một cấu trúc phụ thuộc được chọn, PLS_{predict} sử dụng các giá trị cho các biến chỉ báo của các cấu trúc độc lập của các quan sát trong mẫu giữ lại và áp dụng các giá trị ước lượng của mô hình từ mẫu huấn luyện để tạo ra các dự đoán của các biến chỉ báo của các cấu trúc phụ thuộc trong mẫu giữ lại (Shmueli và cộng sự (2016)). Một sự khác biệt nhỏ giữa các giá trị thực và các giá trị dự đoán cho thấy mô hình có sức mạnh dự đoán cao, và ngược lại.

Theo Hair và cộng sự (2022: trang 222), để đánh giá sức mạnh dự đoán của mô hình, các nhà nghiên cứu nên sử dụng các thống kê dự đoán sau đây: Sai số trung bình tuyệt đối (**MAE**), hoặc căn bậc hai của tổng sai số trung bình (**RMSE**).

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (21)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (22)$$

Trong đó, y_i là giá trị thực của y cho quan sát i ($i = 1, \dots, n$) và \hat{y}_i là giá trị dự đoán của y cho quan sát i .

Như vậy, **MAE** là cách khác biệt tuyệt đối trung bình giữa các quan sát thực tế và dự đoán (trong mẫu thử nghiệm), trong đó tất cả các khác biệt riêng lẻ có trọng số bằng nhau. Một đặc điểm quan trọng của thống kê MAE là các giá trị của nó phụ thuộc vào thang đo của các biến quan sát. Thang đo ít điểm (như Likert 7 điểm) thường có MAE cao hơn so với thang đo nhiều điểm (như thang đo 100 điểm). Tương tự, các giá trị của RMSE cũng phụ thuộc vào thang đo của các biến quan sát. Tuy nhiên, do RMSE lấy bình phương các sai số trước khi tính trung bình, nên nó gán các trọng số lớn hơn cho

các sai số lớn hơn. Điều này thật sự hữu ích khi dữ liệu có nhiều sai số lớn không mong muốn. Chính vì thế, RMSE thường là lựa chọn mặc định trong các phần mềm mô hình hóa với mục đích dự đoán.

Cách giải thích các kết quả như thế nào?

MAE và RMSE của mô hình nào có giá trị nhỏ hơn chứng tỏ mô hình đó tốt hơn. Thông thường ta sẽ so sánh giữa mô hình đường dẫn PLS (hay PLS-SEM) với một mô hình chuẩn là hồi quy tuyến tính (**LM**: linear regression model benchmark). Đối với thống kê **$Q^2predict$** , mô hình chuẩn để so sánh là một mô hình thô (**naïve** benchmark), trong đó chỉ sử dụng các mức trung bình đơn giản của các biến chỉ báo trong mẫu huấn luyện. Theo [Hair và cộng sự \(2022: trang 222\)](#), thống kê $Q^2predict$ (của mô hình thô) dương cho biết sai số dự đoán của mô hình đường dẫn PLS nhỏ hơn sai số dự đoán của mô hình thô. Nói cách khác, $Q^2predict$ dương chứng tỏ mô hình đường dẫn PLS thể hiện sức mạnh dự báo tốt hơn mô hình thô.

Một khi đã xác định $Q^2predict > 0$, ta tiếp tục so sánh các giá trị RMSE (hoặc MAE) của mô hình PLS-SEM với mô hình hồi quy tuyến tính (LM). Có các khả năng sau đây:

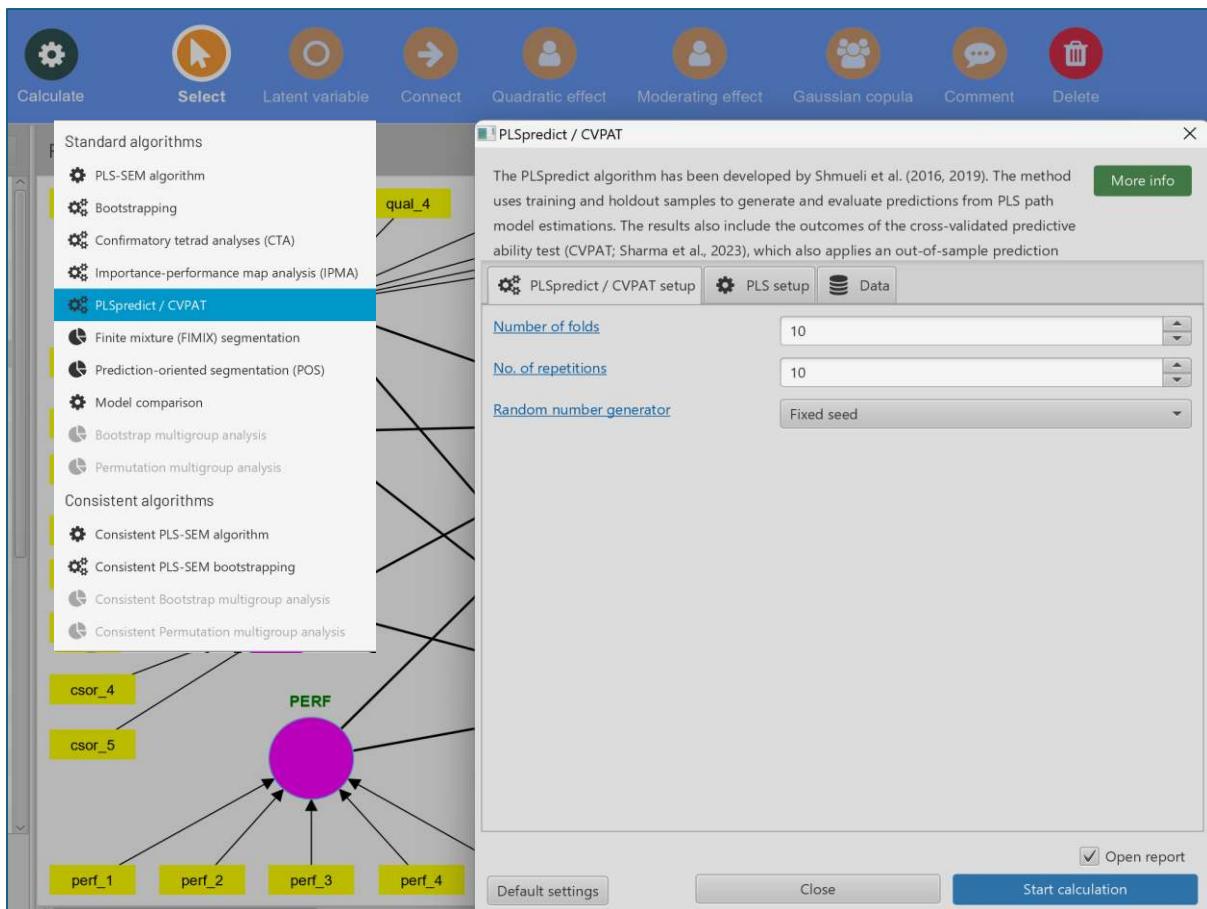
- Nếu tất cả (hoặc phần lớn) các biến chỉ báo trong mô hình PLS-SEM có các giá trị RMSE (hoặc MAE) thấp hơn so với mô hình LM, thì mô hình PLS-SEM có sức mạnh dự đoán cao.
- Nếu số lượng bằng nhau các biến chỉ báo trong mô hình PLS-SEM có các sai số dự đoán (prediction errors) thấp hơn so với mô hình LM, thì mô hình PLS-SEM có sức mạnh dự đoán trung bình.
- Nếu phần nhỏ các biến chỉ báo trong mô hình PLS-SEM có các sai số dự đoán (prediction errors) thấp hơn so với mô hình LM, thì mô hình PLS-SEM có sức mạnh dự đoán thấp.
- Nếu không có biến chỉ báo nào có các giá trị RMSE (hoặc MAE) thấp hơn so với mô hình LM, thì mô hình PLS-SEM không có sức mạnh dự đoán.

Thực hiện PLS_{predict} trên SmartPLS 4 như **Hình 50: Calculate → PLS_{predict} / CVPAT → Start calculation**.

Sau khi thực hiện dự đoán với PLSpredict, ta có thể chọn một trong hai lựa chọn xem kết quả sau đây (hoặc cả hai):

- [Final results → MV prediction summary → Overview => Các biến chỉ báo trung gian \(MV\) => Phổ biến hơn.](#)
- [Final results → LV prediction summary → Overview => Các biến tiềm ẩn \(MV\).](#)

Kết quả dự đoán ở **Hình 51**. Lưu ý là việc phân tích dự báo tập trung vào cấu trúc mục tiêu (target construct) của mô hình (tức cấu trúc CUSL). Thống kê **$Q^2_predict$** của các biến chỉ báo cho cấu trúc CUSL (đó là *cusl_1*, *cusl_2*, và *cusl_3*) đều lớn hơn 0; điều này nói lên rằng mô hình PLS-SEM thể hiện tốt hơn so với mô hình thô.



Hình 50. Các bước thực hiện PLS_{predict} trên SmartPLS 4.

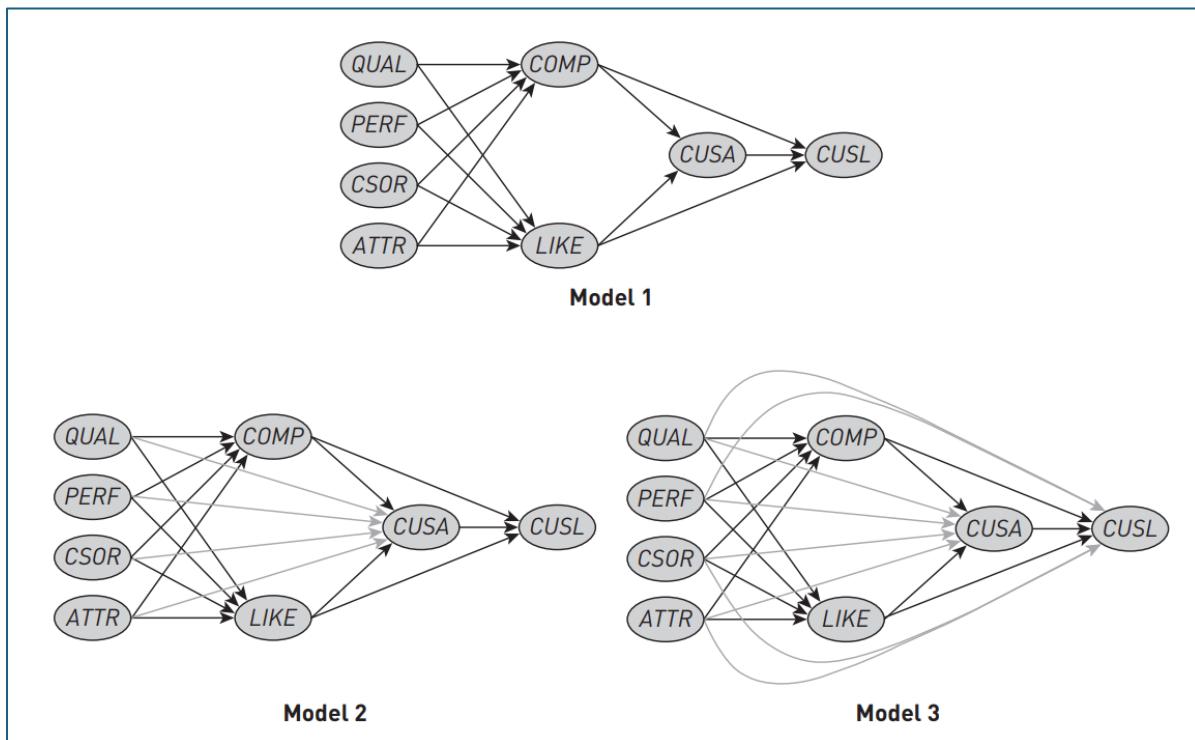
PLSpredict / CVPAT		MV prediction summary - Overview					
			Q ² predict	PLS-SEM_RMSE	PLS-SEM_MAE	LM_RMSE	LM_MAE
comp_1	Graphical	0.473	1.043	0.798	1.070	0.815	
comp_2	Final results	0.364	1.099	0.881	1.110	0.872	
comp_3	MV prediction summary	0.398	1.135	0.898	1.133	0.869	
cusa	Overview	0.245	1.022	0.791	1.019	0.796	
cusl_1	PLS-SEM prediction error (descriptiv	0.259	1.300	0.989	1.310	0.964	
cusl_2	PLS-SEM predictions (descriptives)	0.233	1.523	1.174	1.542	1.170	
cusl_3	LV prediction summary	0.141	1.531	1.155	1.569	1.178	
like_1	CVPAT	0.472	1.128	0.857	1.154	0.881	
like_2	MV predictions and errors	0.360	1.483	1.149	1.546	1.186	
like_3	LV predictions and errors	0.351	1.510	1.170	1.587	1.240	

Hình 51. Kết quả các thống kê dự đoán của mô hình.

Đa số các giá trị RMSE của mô hình PLS-SEM thấp hơn so với mô hình LM (các giá trị màu xanh: sai số PLS-SEM < sai số LM; màu đỏ: sai số PLS-SEM > sai số LM) chứng tỏ mô hình PLS-SEM có sức mạnh dự đoán cao.

Bước 5: So sánh các mô hình (Model comparisons).

Mô hình mở rộng mà ta vừa thực hiện ở những phần trên (như **Model 1** trong **Hình 52**) chỉ là một trong nhiều khả năng có thể xảy ra (nghiên cứu khám phá). Nếu như ta xây dựng mô hình này dựa trên một nền tảng lý thuyết vững chắc thì có thể các kết quả của nó cho các mục đích nghiên cứu hoặc ra quyết định. Tuy nhiên, với một nghiên cứu khám phá thì ta nên ước lượng tất cả các khả năng có thể xảy ra để lựa chọn một mô hình sau cùng tốt nhất. Ngoài **Model 1**, ta có thể xem xét thêm **Model 2** (các cấu trúc nguyên nhân có khả năng ảnh hưởng trực tiếp lên CUSA) và **Model 3** (các cấu trúc nguyên nhân có khả năng ảnh hưởng trực tiếp lên CUSA và CUSL).



Hình 52. Các mô hình khác nhau về lòng trung thành của khách hàng.

Ngoài R^2 thì tiêu chí thường được sử dụng để so sánh giữa các mô là tiêu chí thông tin Bayes (**BIC**). Ta không trình bày các công thức tính toán ở đây. Mô hình nào có giá trị BIC nhỏ hơn thì mô hình đó được coi là tốt hơn. Đối với R^2 thì mô hình nào có giá trị R^2 cao hơn thì mô hình đó được xem là tốt hơn. Tuy nhiên, do R^2 có nhược điểm là thiên vị các mô hình phức tạp với nhiều cấu trúc ngoại sinh hơn, bất kể cấu trúc đó có thực sự liên quan đến cấu trúc nội sinh hay không.

Để thực hiện việc so sánh các mô hình trong **Hình 52**, ta cần phải thực hiện các công việc sau đây trên SmartPLS 4:

- Từ **Extended model** (mô hình có sẵn trong dự án PLS-SEM2024), ta sao chép và đặt tên cho ba mô hình mới là Model 1, Model 2, Model 3; trong đó, Model được giữ nguyên vì hoàn toàn giống với Extended model.
- Mở **Model 2**, chọn “**Connect**” và kết nối các cấu trúc như trong **Hình 52**.

- Mở **Model 3**, chọn “**Connect**” và kết nối các cấu trúc như trong **Hình 52**.
- Uớc lượng Model 1, Model 2, và Model 3, và lưu kết quả của từng mô hình (nhớ đặt tên cho các tệp kết quả).
- Vào từng kết quả ước lượng chọn: **Quality criteria → Model selection criteria**. Ta có thể nhấp vào biểu tượng “**Compare**” để so sánh với một mô hình khác (lưu ý là chỉ một mô hình khác).
- Nên tập trung vào giá trị BIC của biến mục tiêu (CUSL) để so sánh.

Kết quả so sánh được trình bày ở **Bảng 13**.

Bảng 13. So sánh các mô hình theo tiêu chí BIC.

Biến nội sinh	Mô hình 1	Mô hình 2	Mô hình 3
COMP	-314.62	-304.19	-297.58
CUSA	-102.20	-93.97	-97.40
CUSL	-261.60	-261.60	-245.04
LIKE	-252.38	-255.99	-256.67

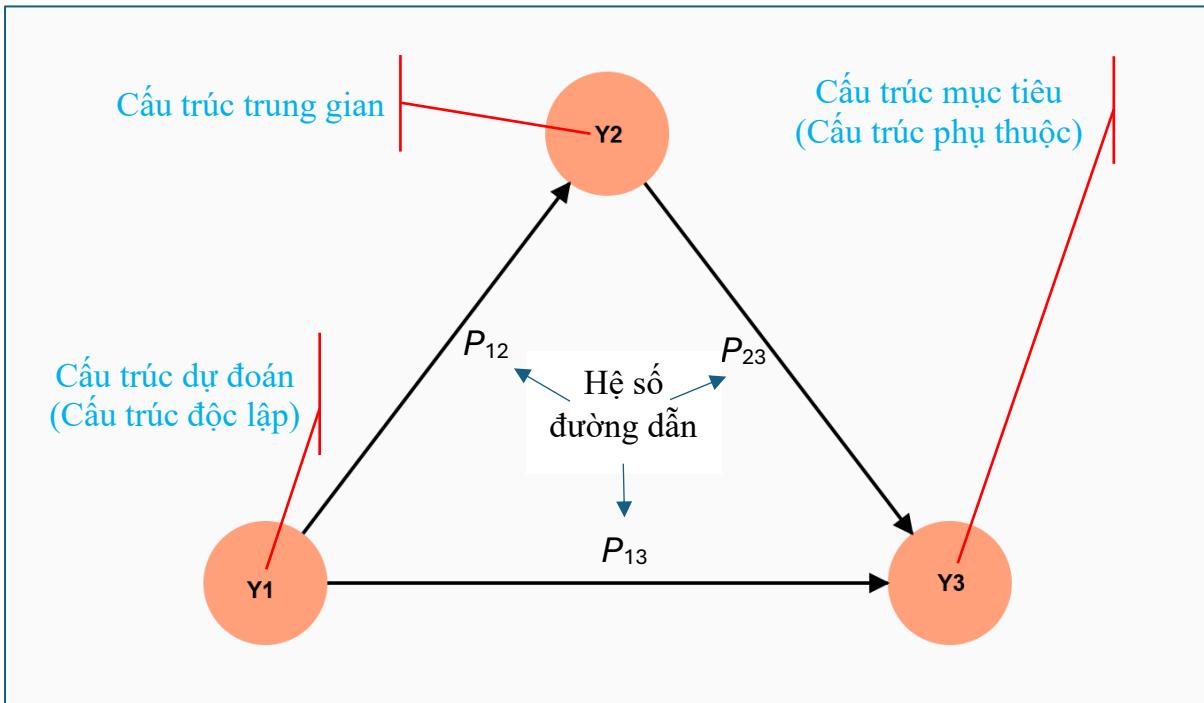
Như vậy, với cấu trúc mục tiêu là CUSL thì Model 1 và Model 2 tốt hơn Model 3. Nếu thêm hai cấu trúc COMP và CUSA thì Model 1 tốt hơn cả Model 2 và Model 3. Riêng cấu trúc LIKE thì Model 2 tốt hơn Model 1 và Model 3. Như vậy, mô hình tốt nhất sẽ là mô hình 1 (nhớ rằng “mô hình càng đơn giản càng tốt”). Ta cũng có thể kiểm tra xem các hệ số đường dẫn của các cấu trúc nguyên nhân có ý nghĩa thống kê hay không. Ta nhận thấy chỉ các cấu trúc QUAL có ảnh hưởng có ý nghĩa thống kê đối với CUSA và CUSL. Như vậy, ta có thể chỉ bổ sung đường dẫn từ QUAL đến CUSA và CUSL và tạo thành một mô hình mới (sau đó tiếp tục so sánh với mô hình 1).

7. BIÊN TRUNG GIAN VÀ BIÊN ĐIỀU TIẾP

7.1 Phân tích ảnh hưởng trung gian/gián tiếp

Tác động trung gian (**mediation**) xảy ra khi một cấu trúc trung gian (**mediator construct**) can thiệp vào quan hệ trực tiếp của các cấu trúc khác. Một sự thay đổi trong một cấu trúc ngoại sinh dẫn đến một sự thay đổi của cấu trúc trung gian, và đến lược nó sự thay đổi của cấu trúc trung gian làm thay đổi cấu trúc nội sinh trong mô hình cấu trúc. Vì thế, một cấu trúc trung gian có vai trò kiểm soát bản chất mối quan hệ giữa hai cấu trúc (ngoại sinh và nội sinh) trong mô hình PLS-SEM. Dưới hình thức đơn giản nhất, mô hình cấu trúc chỉ có một cấu trúc trung gian, nhưng với các mô hình phức tạp hơn thì có thể bao gồm nhiều cấu trúc trung gian. Việc quyết định có hay không có các cấu trúc trung gian hoặc có bao nhiêu cấu trúc trung gian phải dựa trên nền tảng lý thuyết. Ví dụ, ở mô hình đơn giản về lòng trung thành của khách hàng thì CUSA là một cấu trúc trung gian. Còn ở mô hình mở rộng thì COMP, LIKE, và CUSA là các cấu trúc

trung gian. Như vậy, các cấu trúc trung gian là các cấu trúc vừa nội sinh vừa ngoại sinh trong mô hình cấu trúc. Ta xét ví dụ một hình thức đơn giản nhất như ở **Hình 53**.



Hình 53. Mô hình chỉ có một cấu trúc trung gian (**single mediation analysis**).

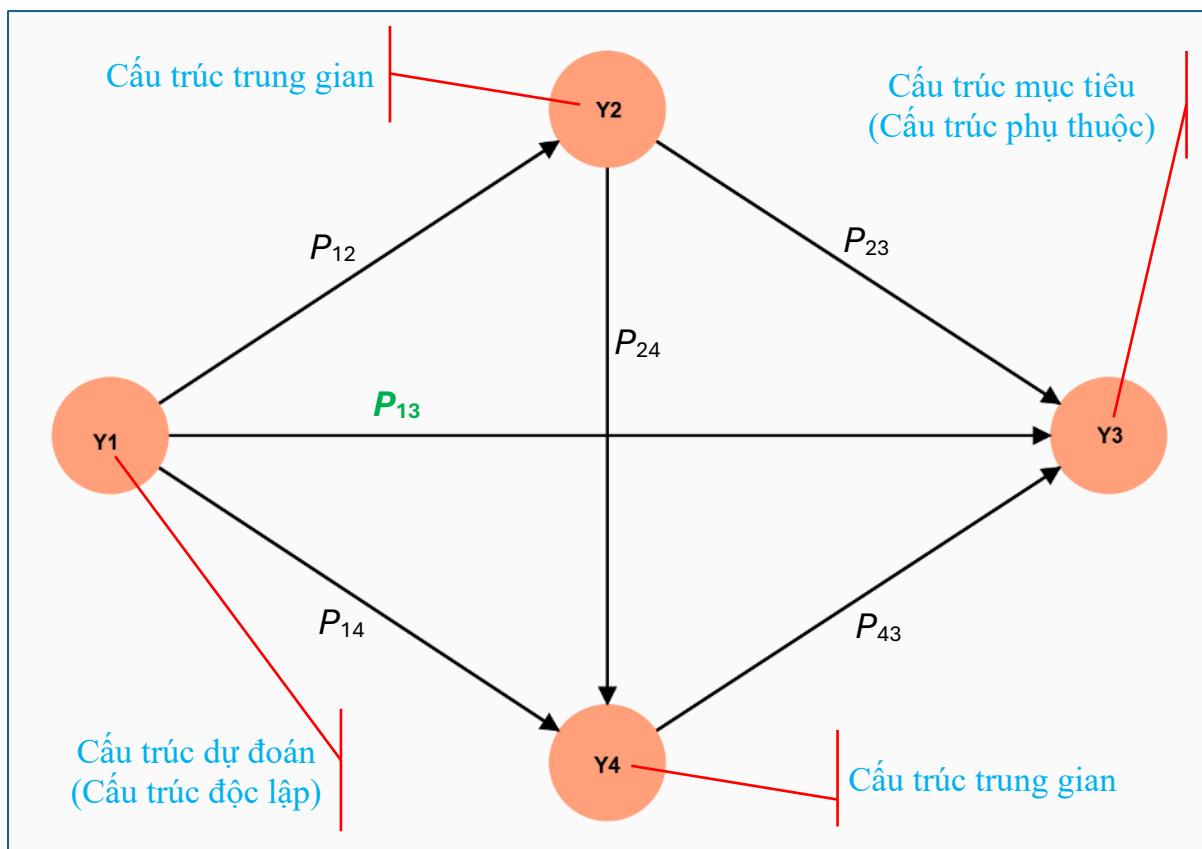
Ảnh hưởng trực tiếp (**direct effect**) của Y_1 lên Y_3 ($Y_1 \rightarrow Y_3$) hệ số p_{13} ; và ảnh hưởng gián tiếp (**indirect effect**) của Y_1 lên Y_3 thông qua Y_2 ($Y_1 \rightarrow Y_2 \rightarrow Y_3$) là tích của hai hệ số $p_{12} \times p_{23}$; và tổng ảnh hưởng của Y_1 lên Y_3 là tổng của hai ảnh hưởng trực tiếp và gián tiếp ($p_{13} + p_{12} \times p_{23}$).

Có các loại ảnh hưởng của cấu trúc Y_1 lên Y_3 như sau:

- Không có ảnh hưởng (**No-effect nonmediation**): Không có ảnh hưởng trực tiếp (p_{13} không có ý nghĩa thống kê) và cũng không có ảnh hưởng gián tiếp ($p_{12} \times p_{23}$ không có ý nghĩa thống kê).
- Chỉ ảnh hưởng trực tiếp (**Direct-only nonmediation**): Có ảnh hưởng trực tiếp (p_{13} có ý nghĩa thống kê), nhưng không có ảnh hưởng gián tiếp ($p_{12} \times p_{23}$ không có ý nghĩa thống kê).
- Ảnh hưởng trung gian bổ sung (**Complementary mediation**): Có cả ảnh hưởng trực tiếp và gián tiếp (p_{13} và $p_{12} \times p_{23}$ đều có ý nghĩa thống kê) và cùng xu hướng (nghĩa là dấu của các hệ số p_{13} và $p_{12} \times p_{23}$ giống nhau).
- Ảnh hưởng trung gian cạnh tranh (**Competitive mediation**): Có cả ảnh hưởng trực tiếp và gián tiếp (p_{13} và $p_{12} \times p_{23}$ đều có ý nghĩa thống kê) nhưng khác xu hướng (nghĩa là dấu của các hệ số p_{13} và $p_{12} \times p_{23}$ khác nhau nhau). Lưu ý: Nếu độ lớn hai tác động bằng nhau thì tổng ảnh hưởng bằng 0.
- Chỉ ảnh hưởng gián tiếp (**Indirect-only mediation**): Có ảnh hưởng gián tiếp ($p_{12} \times p_{23}$ có ý nghĩa thống kê) nhưng không có ảnh hưởng trực tiếp (p_{13} không có ý

nghĩa thống kê). Loại này còn được gọi là ảnh hưởng trung gian toàn phần/hoàn toàn (**full mediation**).

Với một mô hình phức tạp với nhiều cấu trúc trung gian (**multiple mediation analysis**) như **Hình 54** thì ta có thể các khái niệm ảnh hưởng gián tiếp cụ thể (**specific indirect effect**) và tổng ảnh hưởng gián tiếp (**total indirect effect**). Ví dụ, ảnh hưởng gián tiếp cụ thể của Y_1 lên Y_3 thông qua cấu trúc trung gian Y_2 được tính bằng $p_{12} \times p_{23}$; ảnh hưởng gián tiếp cụ thể của Y_1 lên Y_3 thông qua cấu trúc trung gian Y_4 được tính bằng $p_{14} \times p_{43}$; ảnh hưởng gián tiếp cụ thể của Y_1 lên Y_3 thông qua cả cấu trúc trung gian Y_2 và Y_4 được tính bằng $p_{12} \times p_{24} \times p_{43}$. Như vậy, tổng ảnh hưởng gián tiếp của Y_1 lên Y_3 sẽ là: $(p_{12} \times p_{23} + p_{14} \times p_{43} + p_{12} \times p_{24} \times p_{43})$. Tương tự như mô hình chỉ một cấu trúc trung gian thì ảnh hưởng trực tiếp vẫn là p_{13} .



Hình 53. Mô hình nhiều cấu trúc trung gian (**multiple mediation analysis**).

Để biến các ảnh hưởng trung gian có ý nghĩa thống kê hay không ta thường dùng kiểm định Sobel (**Sobel test**). Kiểm định này so sánh mối quan hệ trực tiếp giữa hai cấu trúc độc lập và phụ thuộc với mối quan hệ gián tiếp giữa hai cấu trúc độc lập và phụ thuộc bao gồm cả cấu trúc trung gian (**Helm và cộng sự, 2010**). Kiểm định Sobel giả định phân phối chuẩn của các hệ số ước lượng (p_{12}, p_{23}), và điều này không phù hợp với phương pháp ước lượng phi tham số như PLS-SEM. Hơn nữa, các giả định phân phối chuẩn của kiểm định này thường không thỏa mãn bởi vì tích của hai hệ số $p_{12} \times p_{23}$ không theo phân phối chuẩn. Hơn nữa, kiểm định Sobel yêu cầu các hệ số đường dẫn

không chuẩn hóa (**unstandardized path coefficients**) và thiếu sức mạnh thống kê khi áp dụng cho các cỡ mẫu nhỏ ([Hair và cộng sự, 2022: trang 236](#)). Chính vì thế, các nhà nghiên cứu thường không dùng kiểm định Sobel để đánh giá các kết quả phân tích ảnh hưởng trung gian ([Preacher & Hayes, 2008; Sattler và cộng sự, 2010](#)), đặc biệt là khi mô hình bao gồm các biến tiềm ẩn như PLS-SEM ([Hair và cộng sự, 2022: trang 236](#)).

Thay vì sử dụng kiểm định Sobel, các nhà nghiên cứu sử dụng phương pháp lặp đi lặp lại (**bootstrap**) phân phối mẫu của ảnh hưởng gián tiếp. Cách tiếp cận này được đề xuất bởi [Preacher và Hayes \(2004, 2008\)](#). Bootstrapping không có yêu cầu về dạng phân phối của các biến (cấu trúc) hoặc phân phối mẫu của các thống kê kiểm định và có thể được áp dụng cho các cỡ mẫu nhỏ. Cách tiếp cận này rất phù hợp với phương pháp PLS-SEM và được sử dụng trong SmartPLS ([Hair và cộng sự, 2022: trang 236](#)).

Sau khi ước lượng mô hình mở rộng về lòng trung thành của khách hàng theo phương pháp Bootstrapping, ta có kết quả ảnh hưởng gián tiếp như **Hình 54**.

The screenshot displays the SmartPLS software interface. On the left, there's a tree-like navigation menu with sections like 'Graphical', 'Final results', 'Path coefficients', and 'Total indirect effects'. Under 'Total indirect effects', there are three options: 'Mean, STDEV, T values, p values', 'Confidence intervals', and 'Confidence intervals bias corrected'. An orange arrow points from the 'Total effects' section in the sidebar to a blue callout box containing the Vietnamese text 'Ta làm tương tự với tổng ảnh hưởng'. The main right panel shows a table titled 'Total indirect effects - Mean, STDEV, T values, p values' with various path coefficients and their statistics.

	Original sample (O)	Sample mean (M)	Standard deviation (STDEV)	T statistics (O/STDEV)	P values
ATTR-> CUSA	0.085	0.083	0.030	2.811	0.002
ATTR-> CUSL	0.101	0.099	0.037	2.721	0.003
COMP-> CUSL	0.074	0.074	0.035	2.084	0.019
CSOR-> CUSA	0.086	0.087	0.028	3.127	0.001
CSOR-> CUSL	0.105	0.106	0.034	3.092	0.001
LIKE-> CUSL	0.220	0.220	0.035	6.347	0.000
PERF-> CUSA	0.094	0.096	0.039	2.427	0.008
PERF-> CUSL	0.089	0.091	0.044	2.009	0.022
QUAL-> CUSA	0.228	0.231	0.037	6.175	0.000
QUAL-> CUSL	0.248	0.253	0.043	5.821	0.000

Hình 54. Ảnh hưởng gián tiếp và kiểm định thống kê.

7.2 Phân tích ảnh hưởng điều tiết

Tác động điều tiết (**moderation**) là một biến thứ ba có khả năng làm thay đổi một mối quan hệ cấu trúc trong mô hình. Độ lớn và thậm chí là chiều hướng của một mối quan hệ giữa hai cấu trúc phụ thuộc vào biến thứ ba (gọi là biến điều tiết, **moderator**). Các biến điều tiết này thường là các đặc điểm kinh tế xã hội của đối tượng như giới tính, thu nhập, học vấn, v.v. Ví dụ, các nghiên cứu trước cho thấy mối quan hệ giữa sự hài lòng của khách hàng và lòng trung thành của khách hàng khác nhau theo mức thu nhập (hoặc tuổi, v.v., [Homburg & Giering, 2001](#)). Cụ thể, người có thu nhập cao hơn thì mối quan hệ giữa sự hài lòng và lòng trung thành yếu hơn. Lưu ý, các mối quan hệ điều tiết phải được giả thiết dựa trên cơ sở lý thuyết, và cần được kiểm định một cách cụ thể. Và cũng nên giả thiết là chỉ một mối quan hệ cụ thể trong mô hình hay là tất cả các mối quan hệ trong mô hình.

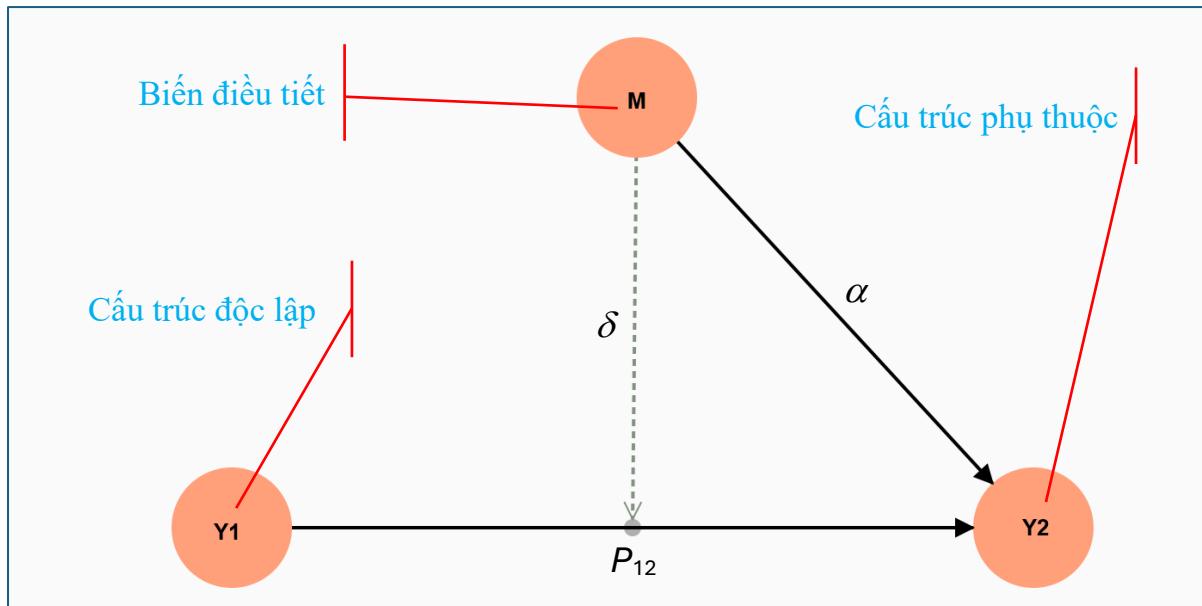
Phân loại biến điều tiết

Biến điều tiết có thể được thể hiện trong các mô hình cấu trúc dưới nhiều hình thức khác nhau. Chúng có thể là các trạng thái có thể quan sát được như giới tính, tuổi, hoặc thu nhập. Nhưng chúng cũng thể hiện dưới dạng những trạng thái không thể quan sát được như mức chấp nhận rủi ro (risk tolerance), thái độ đối với thương hiệu, hoặc sự tham gia. Các biến điều tiết có thể được đo lường với một thang đo duy nhất (một biến quan sát) hoặc nhiều thang đo (nhiều biến quan sát) và sử dụng cả các biến chỉ báo kết quả và nguyên nhân. Cách phân loại quan trọng nhất là chia thành các biến điều tiết phân nhóm (**categorical moderator variable**) và các biến điều tiết liên tục (**continuous moderator variables**).

- Các biến điều tiết phân nhóm (rồi rạc) thường là các biến giả (nhị phân) được mã hóa các giá trị 1 và 0 như nam và nữ, thành thị và nông thôn, ...; nhưng cũng có thể là những biến với nhiều phân nhóm hơn như nhóm thu nhập (thấp, trung bình, cao), sử dụng dịch vụ (hợp đồng ngắn hạn, dài hạn, thuê bao trả trước), ...
- Các biến điều tiết liên tục là những biến với bản chất là liên tục và có thể quan sát được như tuổi, kinh nghiệm theo số năm làm việc, thu nhập tính bằng triệu đồng, ... hoặc là những biến không thể quan sát trực tiếp mà phải thông qua các biến chỉ báo được đo theo các thang đo Likert, Stapel, ...

Mô hình hóa các ảnh hưởng điều tiết

Để hiểu hơn các ảnh hưởng điều tiết (**moderating effects**) được mô hình hóa như thế nào trong các mô hình PLS-SEM, ta hãy xem ví dụ đơn giản ở **Hình 55**. Mỗi quan hệ này có thể được biểu diễn dưới dạng phương trình như sau (giống như biến tương tác trong các mô hình hồi quy truyền thống):

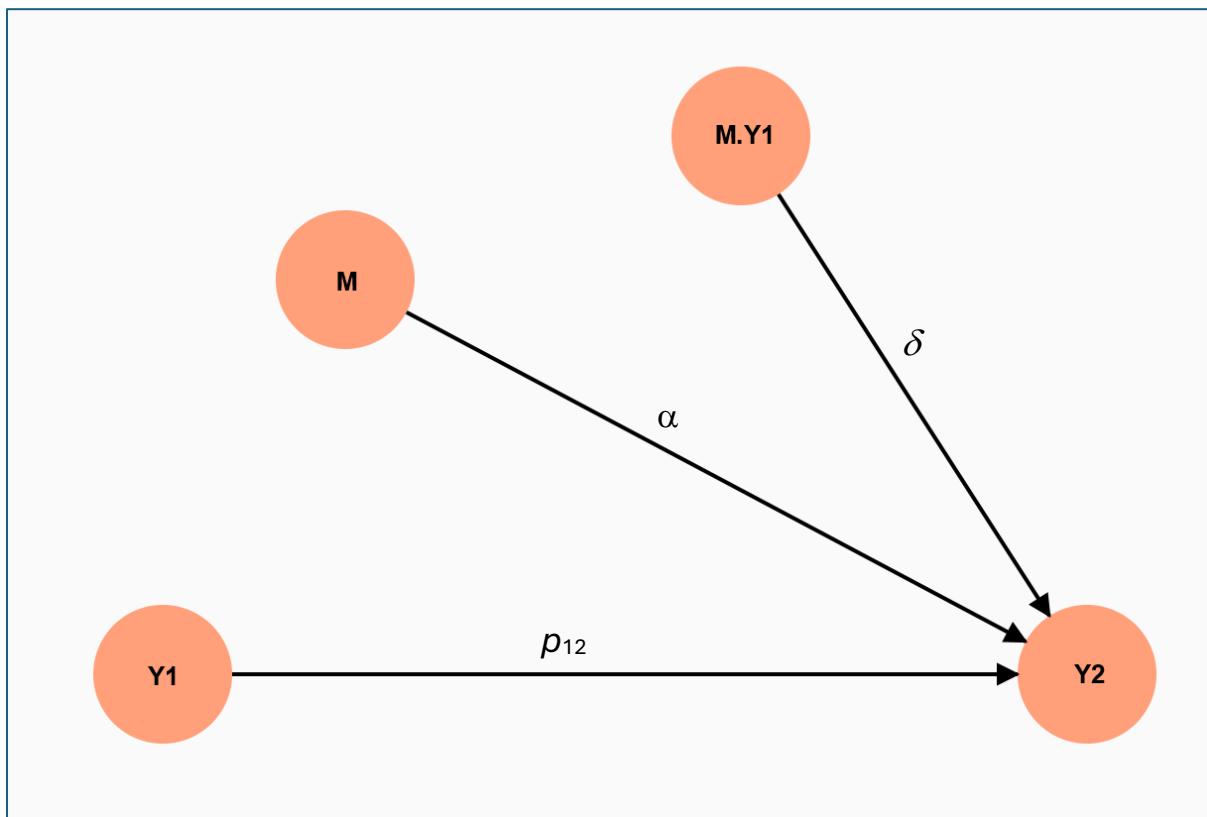


Hình 55. Mô hình điều tiết đơn giản.

$$Y_2 = p_{12}Y_1 + \delta M Y_1 + \alpha M = (p_{12} + \delta M)Y_1 + \alpha M \quad (23)$$

Hệ số δ (nếu có ý nghĩa thống kê) thể hiện ảnh hưởng p_{12} thay đổi như thế nào khi biến điều tiết M tăng lên hay giảm đi một độ lệch chuẩn (với lựa chọn “**Standardized**”) khi thực hiện ước lượng mô hình PLS-SEM. Lưu ý: MY_1 được gọi là biến tương tác (**interaction term**).

Theo phương trình (23) thì mô hình điều tiết đơn giản có thể được biểu diễn theo một cách khác như ở **Hình 56**.

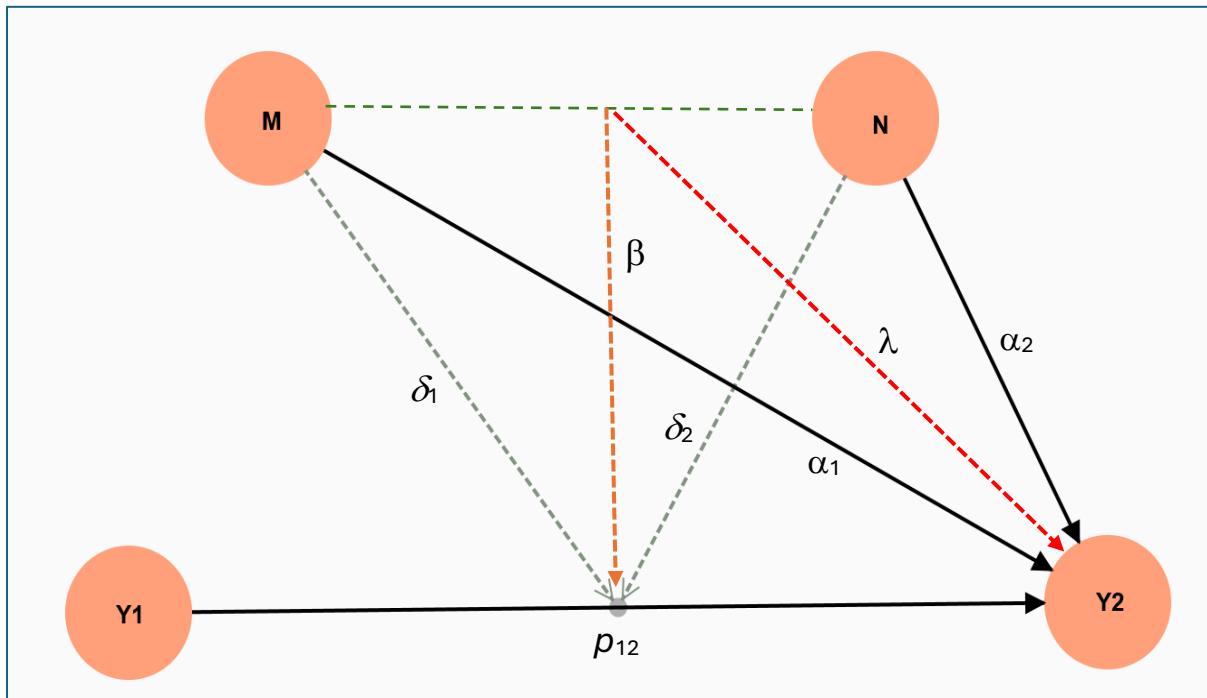


Hình 56. Cấu trúc khác của mô hình điều tiết đơn giản.

Cách mô hình hóa ở trên đơn giản chỉ với một biến điều tiết duy nhất, và được gọi là tương tác hai chiều (**two-way interaction**), bởi vì biến điều tiết M tương tác với một biến khác (cấu trúc ngoại sinh Y_1). Khi thực hiện một mô hình với hơn một biến điều tiết (**multiple moderator analysis**), thì phương trình trở nên phức tạp hơn vì có thể có sự tương tác ba chiều (**three-way interaction**); nghĩa là ngoài sự tương tác với cấu trúc ngoại sinh thì các biến điều tiết còn có thể tương tác với nhau. Ví dụ, bây giờ ta có hai biến tương tác M và N như ở **Hình 57**, thì phương trình thể hiện mối quan hệ này sẽ trở nên phức tạp hơn như sau:

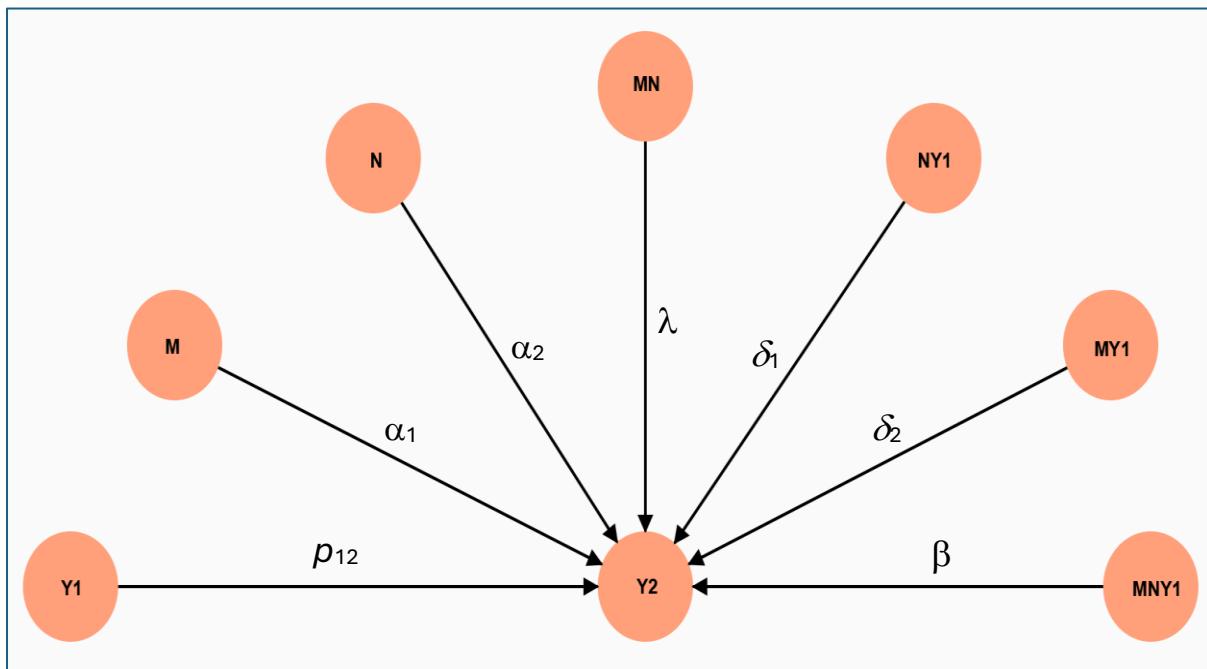
$$\begin{aligned} Y_2 &= p_{12}Y_1 + \delta_1 M Y_1 + \delta_2 N Y_1 + \alpha_1 M + \alpha_2 N + \lambda MN + \beta MNY_1 \\ &= (p_{12} + \delta_1 M + \delta_2 N + \beta MN)Y_1 + \alpha_1 M + \alpha_2 N + \lambda MN \end{aligned} \quad (24)$$

Hệ số δ_1 , δ_2 và β (nếu có ý nghĩa thống kê) thể hiện ảnh hưởng p_{12} thay đổi như thế nào khi các biến điều tiết M và N tăng lên hay giảm đi một độ lệch chuẩn (với lựa chọn “**Standardized**”) khi thực hiện ước lượng mô hình PLS-SEM.



Hình 57. Mô hình điều tiết phức tạp.

Mô hình **Hình 57** có thể được thể hiện lại một cách khác theo phương trình (24) như **Hình 58** sau đây:



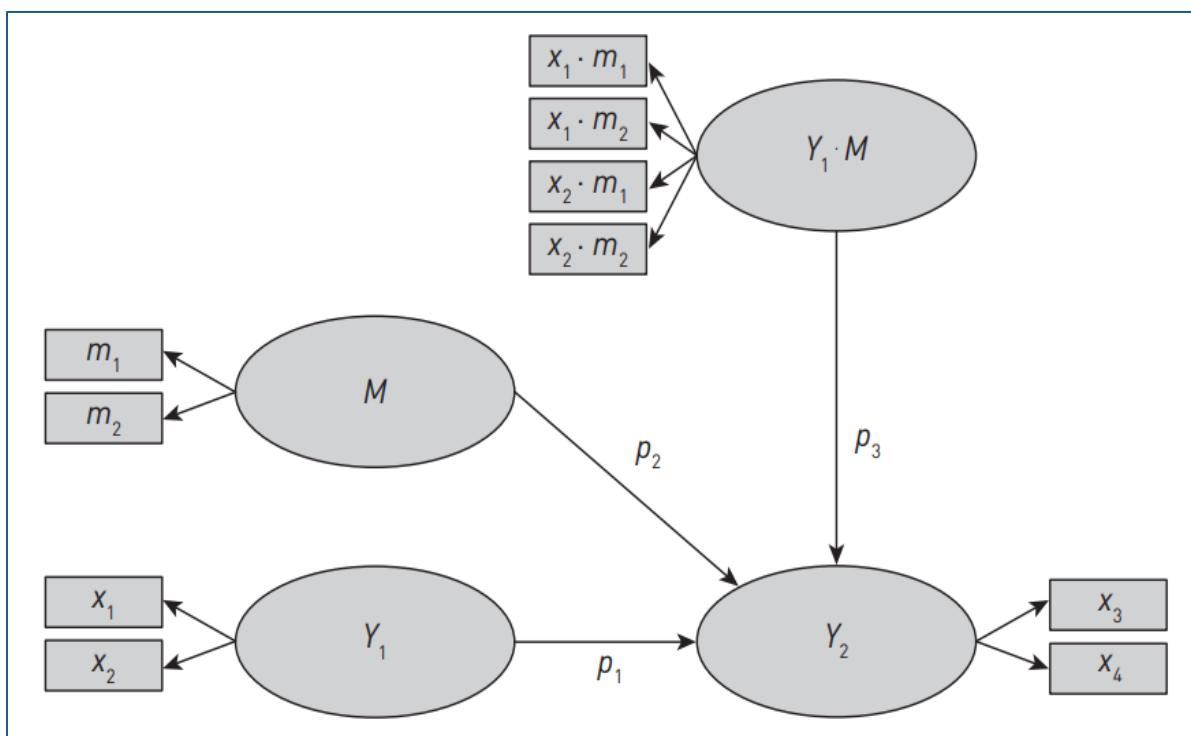
Hình 58. Cấu trúc khác của mô hình điều tiết phức tạp.

Lưu ý rất quan trọng: Chúng ta đang nói biến điều tiết trong sự tương tác với (các) cấu trúc ngoại sinh (tức chỉ đóng vai trò biến độc lập trong mô hình cấu trúc). Ở phần sau sẽ giới thiệu thêm về trường hợp tương tác giữa (các) biến điều tiết và (các) cấu trúc trung gian trong mô hình PLS-SEM.

Tạo các biến tương tác

Cách 1: Cách tiếp cận biến chỉ báo tích (Product indicator approach)

Đây là cách tiếp cận chuẩn để tạo (các) biến tương tác trong các phân tích dựa vào hồi quy và có thể áp dụng cho PLS-SEM. Cách tiếp cận biến chỉ báo tích là việc nhân mỗi biến chỉ báo của cấu trúc ngoại sinh với mỗi biến chỉ báo của biến điều tiết. Nếu biến điều tiết chỉ gồm một biến quan sát duy nhất thì chỉ nhân mỗi biến chỉ báo của cấu trúc ngoại sinh với biến điều tiết ([Chin và cộng sự, 2003](#)). Cách tiếp cận này được minh họa ở **Hình 59**. Cách tiếp cận này có nhược điểm là có thể gây ra sự cộng tuyến trong mô hình đường dẫn PLS ([Hair và cộng sự, 2022: trang 250](#)).



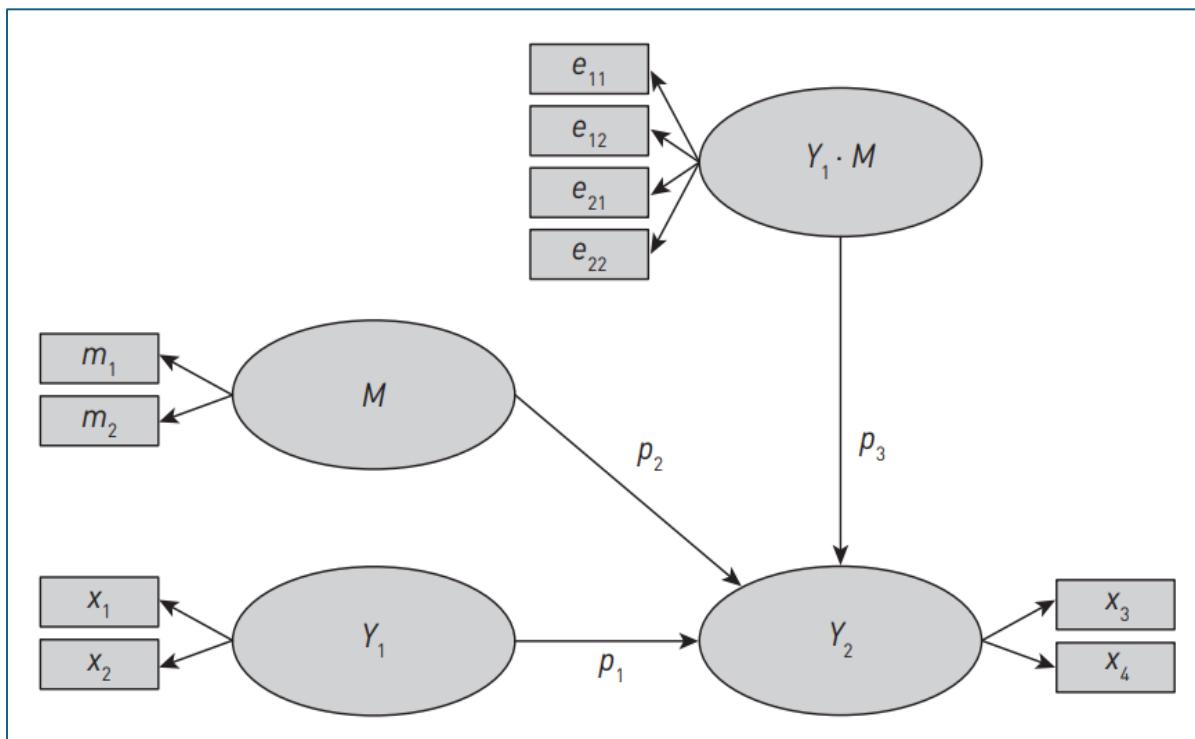
Hình 59. Cách tiếp cận biến chỉ báo tích ([Hair và cộng sự, 2022: trang 249](#)).

Cách 2: Cách tiếp cận trực giao (Orthogonalizing approach)

Đây là một sự mở rộng của cách tiếp cận chỉ báo tích (nhằm giảm khả năng gây ra sự cộng tuyến). Cách tiếp cận này dựa trên cách tiếp cận biến chỉ báo tích và yêu cầu tạo ra tất cả các biến chỉ báo tích của biến tương tác. Với mô hình ở **Hình 60** (biến ngoại sinh Y_1 có 2 biến chỉ báo kết quả, và biến điều tiết M có 2 biến chỉ báo kết quả), ta có thể tạo ra bốn biến chỉ báo tích: $x_1 \cdot m_1$, $x_1 \cdot m_2$, $x_2 \cdot m_1$, và $x_2 \cdot m_2$. Bước tiếp theo hồi quy tùng biến chỉ báo tích mới tạo ra theo tất cả các biến chỉ báo của cấu trúc ngoại sinh

và biến điều tiết. Hệ thống các phương trình hồi quy theo phương pháp OLS (bình phương bé nhất thông thường) như ở phương trình (25). Ta không quan tâm đến các hệ số ước lượng từ các phương trình này; thay vào đó là các phần dư (e_{11}, e_{12}, e_{13} , và e_{14}). Cách tiếp cận trực giao sử dụng các phần dư chuẩn hóa (**standardized residuals**) e như các biến chỉ báo cho các biến tương tác.

$$\begin{aligned}x_1 \cdot m_1 &= b_{1,11} \cdot x_1 + b_{2,11} \cdot x_2 + b_{3,11} \cdot m_1 + b_{4,11} \cdot m_2 + e_{11} \\x_1 \cdot m_2 &= b_{1,12} \cdot x_1 + b_{2,12} \cdot x_2 + b_{3,12} \cdot m_1 + b_{4,12} \cdot m_2 + e_{12} \\x_2 \cdot m_1 &= b_{1,21} \cdot x_1 + b_{2,21} \cdot x_2 + b_{3,21} \cdot m_1 + b_{4,21} \cdot m_2 + e_{21} \\x_2 \cdot m_2 &= b_{1,22} \cdot x_1 + b_{2,22} \cdot x_2 + b_{3,22} \cdot m_1 + b_{4,22} \cdot m_2 + e_{22}\end{aligned}\tag{25}$$



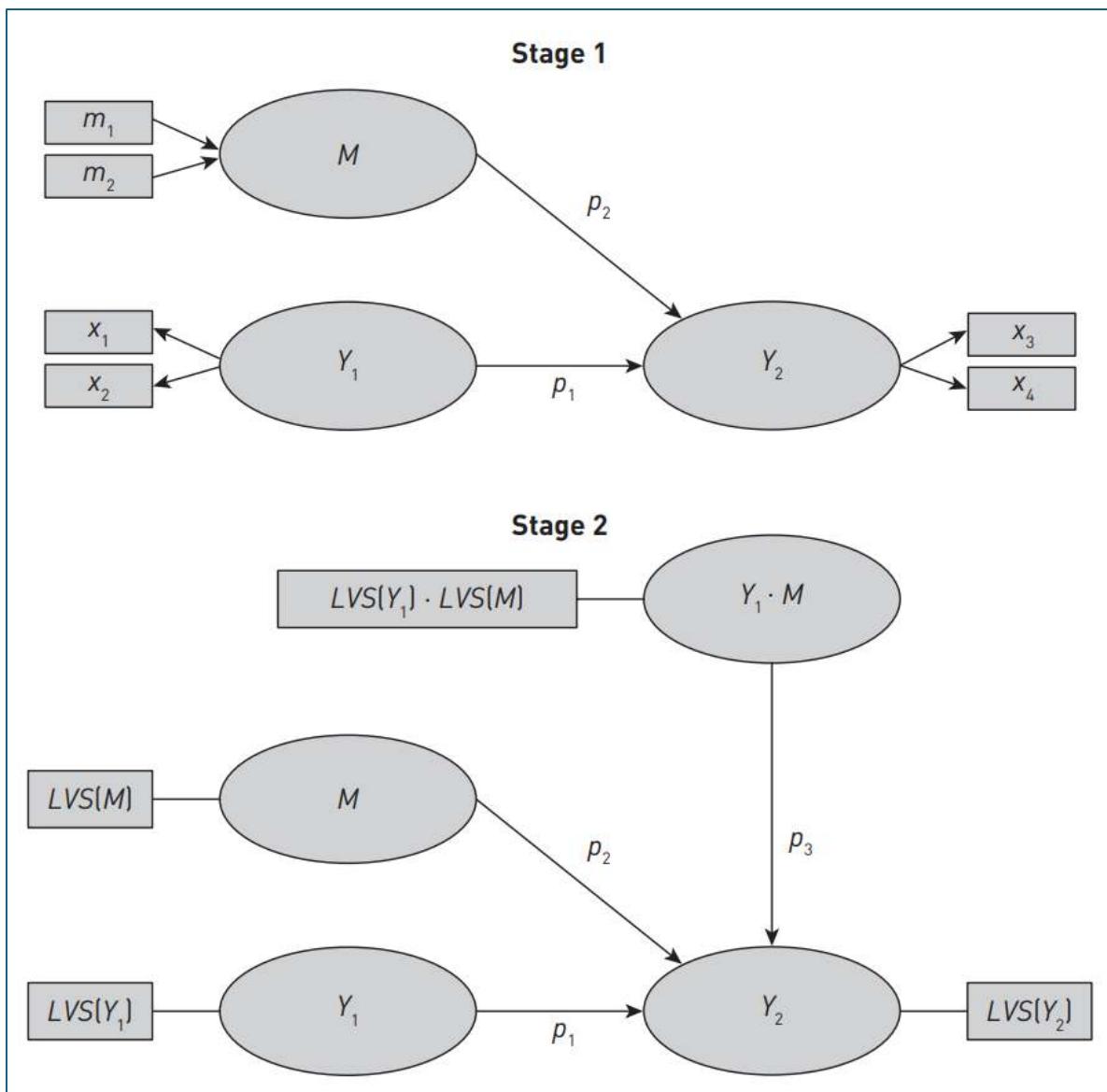
Hình 60. Cách tiếp cận trực giao (Hair và cộng sự, 2022: trang 251).

Cách 3: Cách tiếp cận hai bước (Two-stage approach)

Hai bước trong cách tiếp cận này là:

- **Bước 1:** Mô hình ảnh hưởng chính (main effect model); tức là mô hình không gồm (các) biến tương tác, được ước lượng để thu các điểm số của các biến tiềm ẩn ngoại sinh [**LVS(Y₁)**] và biến điều tiết [**LVS(M)**]. Các biến tiềm ẩn này được lưu lại để phân tích tiếp theo trong bước 2. Trường hợp này ta đang giả sử các biến chỉ báo của **M** là các biến chỉ báo nguyên nhân. Như phần trên của **Hình 61**. Lưu ý: Các biến tiềm ẩn được đặt tên là LVS(M) và LVS(Y₁); với **LVS** là viết tắt của “**Latent Variable Score**”.

- **Bước 2:** Các điểm số của biến tiềm ẩn của cấu trúc ngoại sinh và biến điều tiết từ bước 1 nhận với nhau để tạo ra một thước đo duy nhất (single-item) được sử dụng để đo lường biến tương tác. Như phần dưới của **Hình 61.** Lưu ý: Các điểm số của biến cấu trúc ngoại sinh và biến điều tiết chỉ là một biến riêng lẻ duy nhất nên sẽ không có mũi tên.



Hình 61. Cách tiếp cận hai bước (Hair và cộng sự, 2022: trang 252).

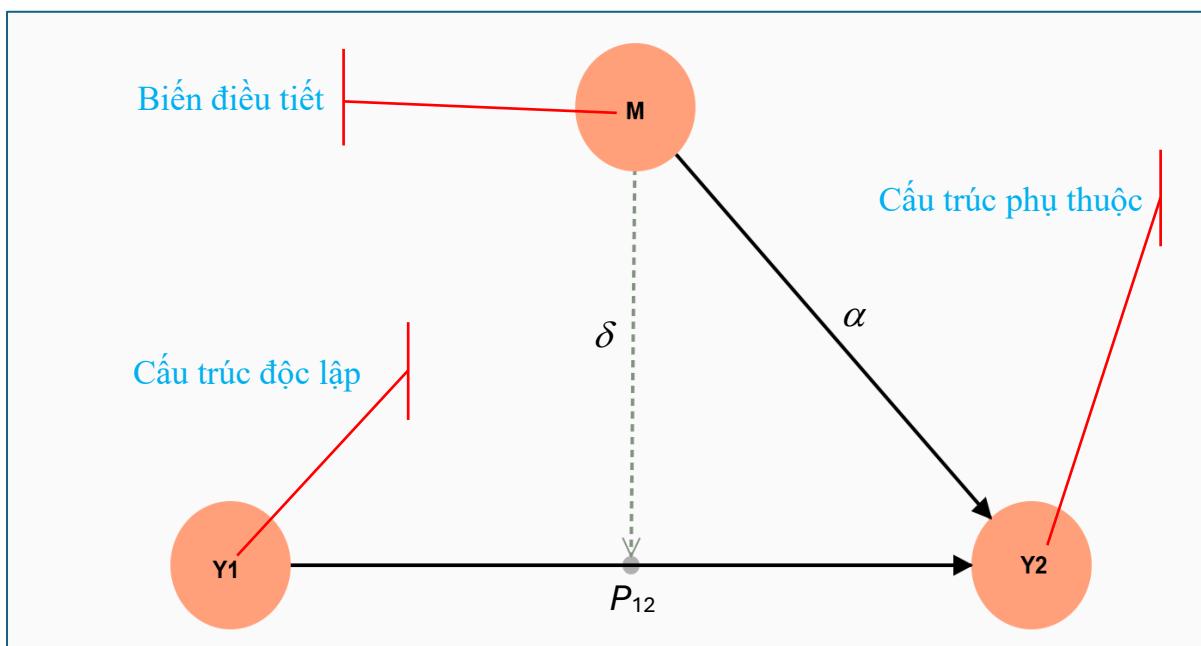
Theo Hair và cộng sự (2022: trang 253), cách tiếp cận biến chỉ báo tích và cách tiếp cận trực giao chỉ có thể áp dụng với các **mô hình đo lường kết quả**. Trái lại, cách tiếp cận hai bước thì có thể áp dụng một cách tổng quát hơn; nghĩa là, bất kể (các) cấu trúc ngoại sinh và (các) cấu trúc điều tiết ở dạng mô hình đo lường kết quả hay mô hình đo lường nguyên nhân. Tùy theo dạng của các biến chỉ báo của biến/cấu trúc điều tiết là gì thì ta lựa chọn cách tiếp cận thích hợp. Vì thế, Hair và cộng sự (2022: trang 253) đề xuất nên sử dụng cách tiếp cận hai bước trong hầu hết các trường hợp.

Đánh giá mô hình

Theo Hair và cộng sự (2022: trang 253), các tiêu chí đánh giá các mô hình đo lường và mô hình cấu trúc cũng áp dụng cho các mô hình có biến điều tiết. Khi đánh giá các mô hình đo lường kết quả (tức là biến/cấu trúc điều tiết hình thành từ các biến chỉ báo kết quả), thì (các) biến điều tiết phải thỏa mãn tất cả các tiêu chí về độ tin cậy nhất quán nội bộ, giá trị hội tụ, và giá trị phân biệt. Tương tự, tất cả các tiêu chí đánh giá mô hình đo lường nguyên nhân như giá trị hội tụ, sự cộng tuyế, ý nghĩa và phù hợp thống kê đều áp dụng cho mô hình có biến điều tiết hình thành từ các biến chỉ báo nguyên nhân. Tuy nhiên, đối với các biến tương tác thì không có các yêu cầu kiểm định như vậy. Sau cùng, ta cũng cần xem xét các tiêu chí để đánh giá mô hình cấu trúc khi có (các) biến điều tiết. Độ lớn ảnh hưởng f^2 là thước đo phổ biến để đánh giá tầm quan trọng của (các) biến tương tác. Nhắc lại rằng, các giá trị độ lớn ảnh hưởng f^2 là 0.02, 0.15, và 0.35 cho biết độ lớn ảnh hưởng của một biến tương tác lần lượt là nhỏ, vừa, và lớn (Cohen, 1988).

Giải thích các kết quả

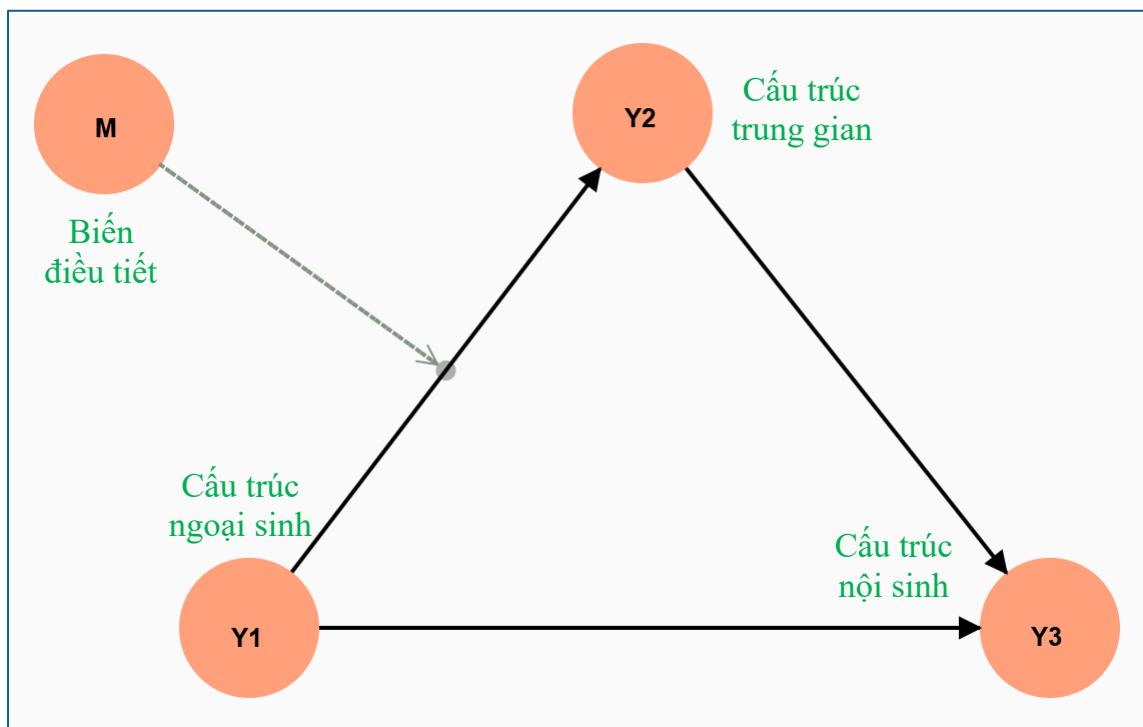
Khi giải thích các kết quả của một phân tích điều tiết, mối quan tâm chính của ta là *sự ý nghĩa của biến tương tác* [tức là biến MY_1 ở phương trình (23); MY_1 , NY_1 và MNY_1 ở phương trình (24)]. Nếu ảnh hưởng của biến tương tác lên cấu trúc ngoại sinh có ý nghĩa thống kê, thì ta kết luận rằng một biến điều tiết M có ảnh hưởng điều tiết có ý nghĩa thống kê lên mối quan hệ giữa Y_1 và Y_2 (như **Hình 55**). Để đánh giá ý nghĩa thống kê của một biến tương tác, ta phải sử dụng kỹ thuật Bootstrapping như đã trình bày ở các phần trước. Nếu như sự điều tiết là có ý nghĩa thống kê (nghĩa khác hệ số của biến tương tác khác 0), bước tiếp theo là ta phải xác định sức mạnh của ảnh hưởng điều tiết (Hair và cộng sự, 2022: trang 253).



Với **Hình 55**, giá trị ước lượng của p_{12} (“nếu p_{12} có ý nghĩa thống kê”) thể hiện sức mạnh/độ lớn của mối quan giữa Y_1 và Y_2 khi hệ số ước lượng của biến tương tác MY_1 bằng 0 (tức $\delta = 0$). Với $\delta \neq 0$, nếu biến điều tiết tăng (hoặc giảm) một độ lệch chuẩn thì ảnh hưởng (trực tiếp, đơn giản¹⁵) p_{12} kỳ vọng thay đổi với giá trị là δ (tùy vào dấu của δ). Nhắc lại là ta thường ước lượng các mô hình PLS-SEM trên SmartPLS 4 với lựa chọn kết quả là chuẩn hóa (**standardized**). Ví dụ, nếu hệ số $p_{12} = 0.3$ và ảnh hưởng điều tiết $\delta = -0.10$, thì ta kỳ vọng mối quan hệ giữa Y_1 và Y_2 giảm từ 0.3 về $0.3 + (-0.10) = 0.2$, nếu giá trị trung bình của biến điều tiết M tăng thêm một đơn vị độ lệch chuẩn (Henseler & Fassott, 2010). Lưu ý, với mô hình có nhiều cấu trúc ngoại sinh, thì ta phải thêm “khi các cấu trúc khác giữ nguyên không đổi” khi giải thích kết quả.

Ảnh hưởng trung gian bị điều tiết và ảnh hưởng điều tiết qua trung gian (Moderated mediation and mediated moderation)

Phần cuối này hơi khó một chút! Tuy nhiên, chúng ta hãy cố gắng để nghe hết toàn bộ câu chuyện. Trong một số trường hợp thì người nghiên cứu có thể xem xét các mối quan hệ hay ảnh hưởng phức tạp hơn, tùy vào các mục tiêu nghiên cứu.



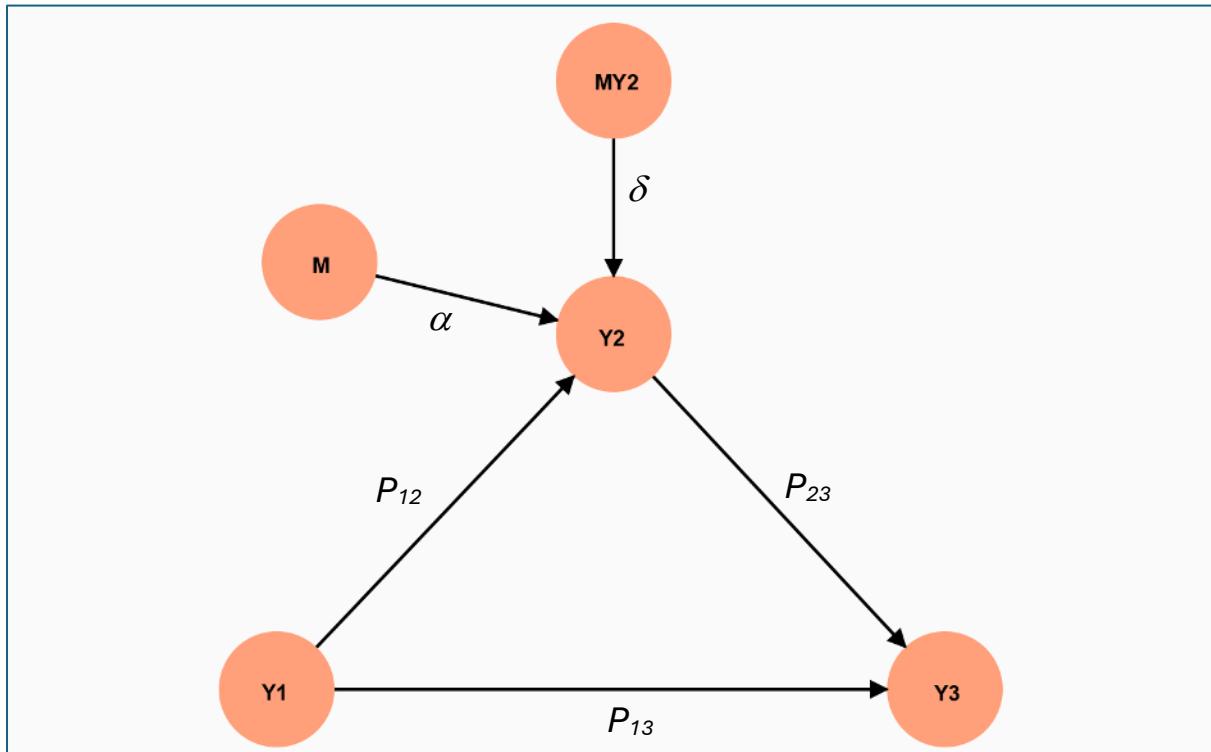
Hình 62. Mô hình ảnh hưởng trung gian bị điều tiết đơn giản.

Ảnh hưởng trung gian bị điều tiết (moderated mediation) xảy ra khi một biến điều tiết tương tác với một cấu trúc trung gian¹⁶ (**mediator construct**); nghĩa là, giá trị của

¹⁵ Ảnh hưởng đơn giản (simple effect) bởi vì ví dụ hiện tại chỉ có một cấu trúc ngoại sinh và một biến điều tiết.

¹⁶ Nhắc lại rằng ở phần trước ta chỉ giới hạn một biến điều tiết tương tác với một cấu trúc ngoại sinh (chỉ đóng vai trò cấu trúc ngoại sinh như **QUAL**, **PERF**, **CSOR**, và **ATTR** trong mô hình mở rộng về lòng trung thành của khách hàng),

ánh hưởng gián tiếp (hay tổng ảnh hưởng gián tiếp) thay đổi phụ thuộc vào giá trị của biến điều tiết. Trường hợp này cũng được gọi là ảnh hưởng gián tiếp có điều kiện (**conditional indirect effect**) bởi vì giá trị của (tổng) ảnh hưởng gián tiếp tùy vào giá trị của biến điều tiết (Hayes, 2015). Mô hình ảnh hưởng gián tiếp bị điều tiết đơn giản được minh họa ở **Hình 62**. Mô hình này được mô hình hóa dưới dạng biến tương tác như được minh họa ở **Hình 63**.



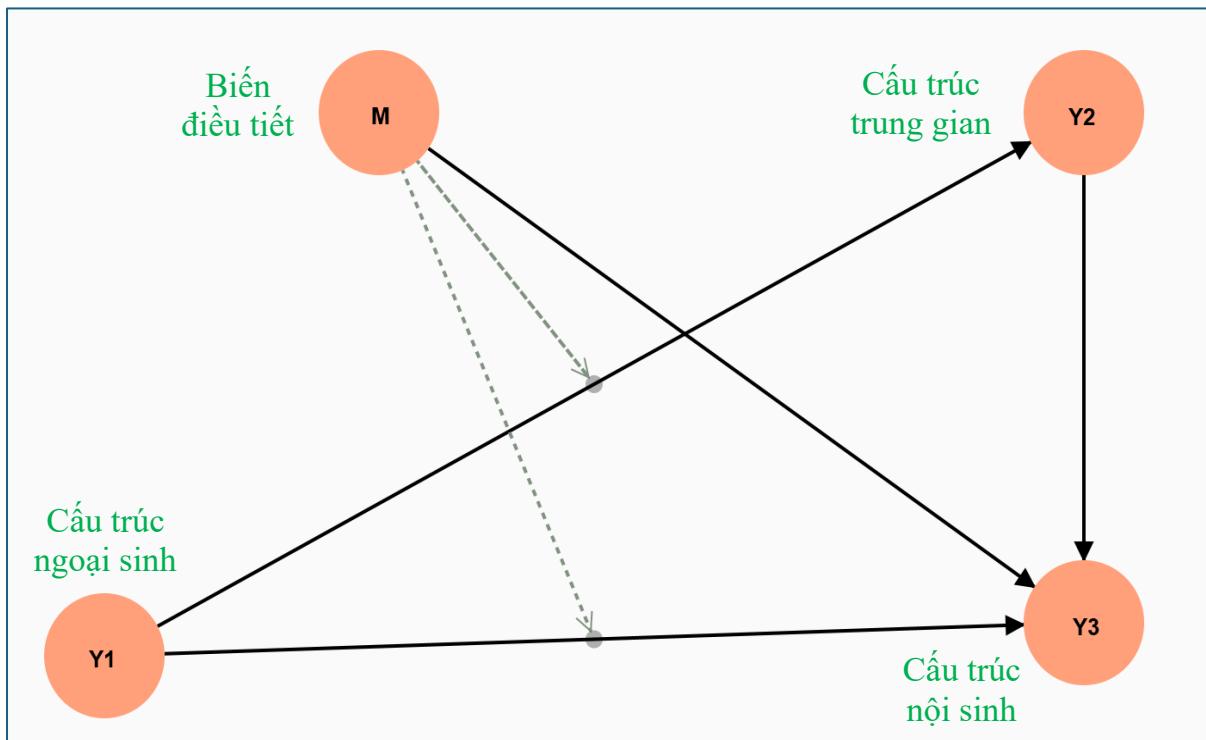
Hình 63. Mô hình hóa mô hình ảnh hưởng trung gian bị điều tiết đơn giản.

Với mô hình đơn giản này, ảnh hưởng gián tiếp có điều kiện của Y_1 lên Y_3 được thể hiện qua công thức sau đây: $(p_{12} + \delta \cdot M) \cdot p_{23} = p_{12} \cdot p_{23} + \delta \cdot p_{23} \cdot M$. Theo Muller và cộng sự (2005) và Preacher và cộng sự (2007) thì để đảm bảo một sự điều tiết có ý nghĩa thống kê thì yêu cầu hệ số ước lượng δ phải có ý nghĩa thống kê.

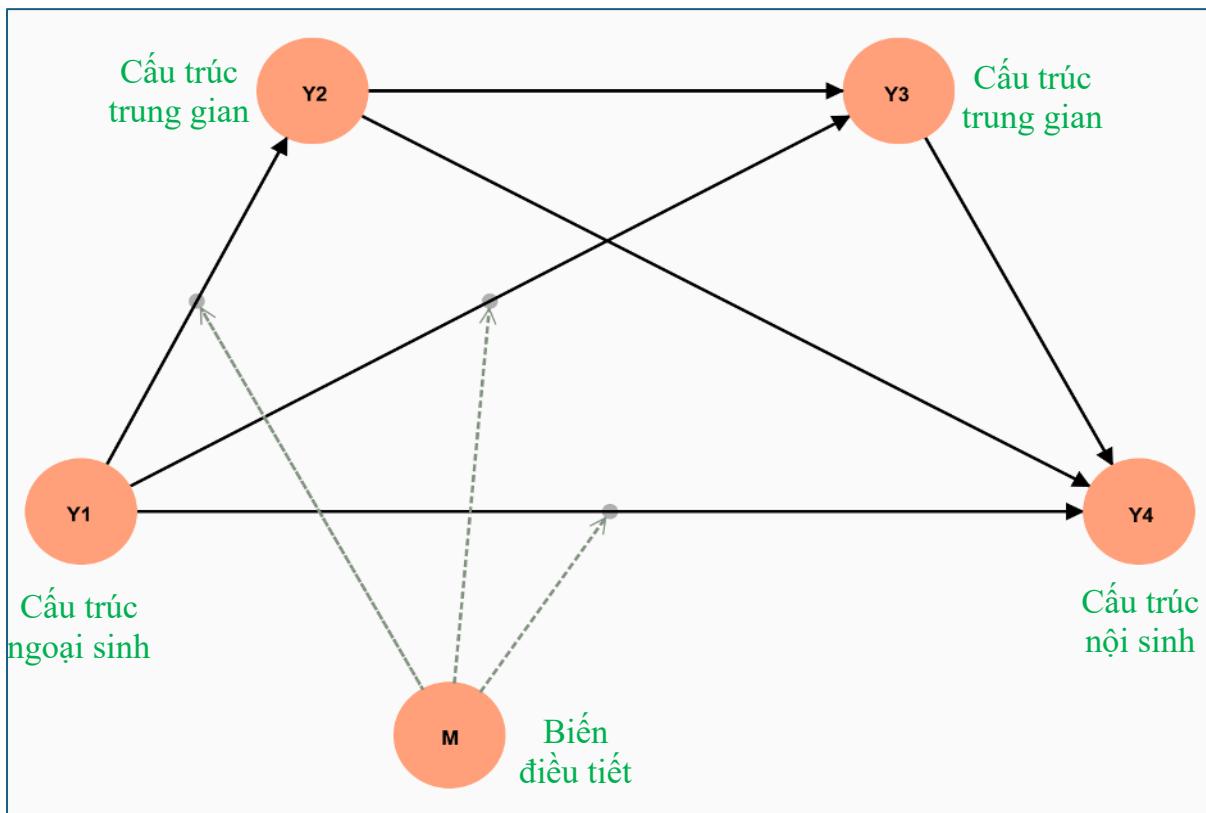
Trong nhiều nghiên cứu ứng dụng, thì biến điều tiết tác động cả lên ảnh hưởng trực tiếp và ảnh hưởng gián tiếp của một cấu trúc ngoại sinh lên cấu trúc nội sinh như các **Hình 64 và 65**. Các mô hình này là mở rộng của mô hình giản đơn ở **Hình 62**.

Sự điều tiết bị ảnh hưởng trung gian (mediated moderation) là trường hợp một cấu trúc trung gian can thiệp vào ảnh hưởng điều tiết theo cách trong đó một sự thay đổi trong biến tương tác dẫn đến một sự thay đổi của cấu trúc trung gian, và thay đổi này dẫn đến sự thay đổi trong cấu trúc phụ thuộc trong mô hình có biến điều tiết. Hayes (2018) có lời khuyên không nên thực hiện kiểm định sự điều tiết bị ảnh hưởng trung gian vì kết quả kiểm định không cung cấp thêm thông tin gì về các ảnh hưởng của mô hình đường dẫn. Hơn nữa, việc kiểm định này thường rất thách thức và không thể thiết

lập nền tảng lý thuyết hợp lý cho các mô hình điều tiết bị ảnh hưởng trung gian. Vì vậy, các nhà nghiên cứu nên tập trung và phân tích ảnh hưởng trung gian bị điều tiết.



Hình 64. Mô hình ảnh hưởng trung gian bị điều tiết phức tạp 1.



Hình 65. Mô hình ảnh hưởng trung gian bị điều tiết phức tạp 2.

Áp dụng phân tích điều tiết với mô hình về lòng trung thành của khách hàng

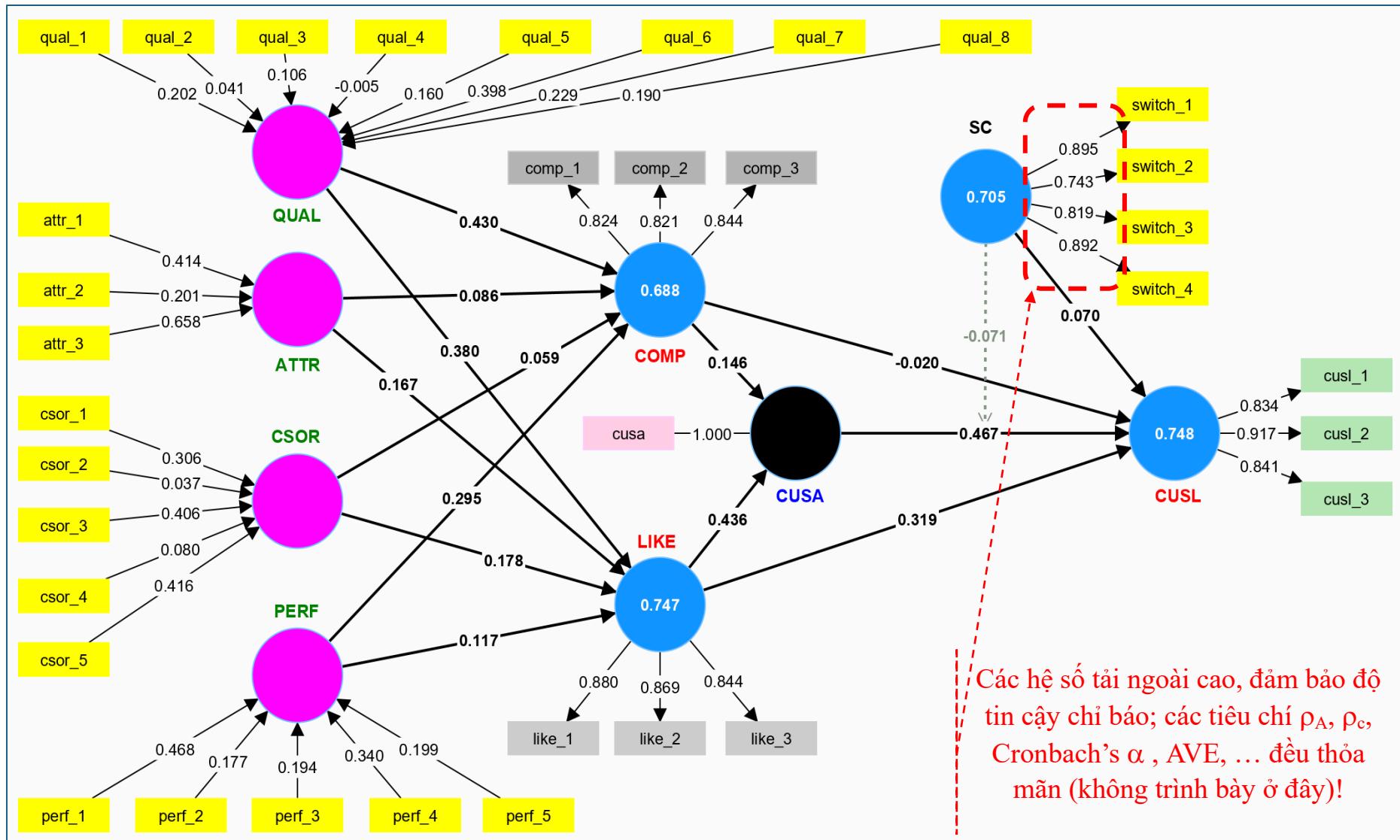
Ta sẽ không thực hành lại việc phân tích ảnh hưởng trung gian bởi vì đã ta đã thực hiện ở các phần trước. Phần này sẽ tập trung phân tích ảnh hưởng điều tiết của ba loại biến điều tiết sau đây:

- Chi phí chuyển đổi (Switching costs: SC) – một dạng cấu trúc kết quả từ các biến chỉ báo kết quả switch_1, switch_2, switch_3, và switch_4 với thang đo Likert 5 điểm;
- Loại dịch vụ (servicetype) – một dạng biến phân nhóm với hai giá trị 0 và 1 (hợp đồng/trả trước);
- Nhà cung cấp (serviceprovider) – Một dạng số liệu phân nhóm với 5 giá trị [1 = Phần mềm máy tính, 2 = Bán lẻ Internet, 3 = Cung cấp dịch vụ Internet, 4 = Máy tính cá nhân, 5 = Dịch vụ video trực tuyến].

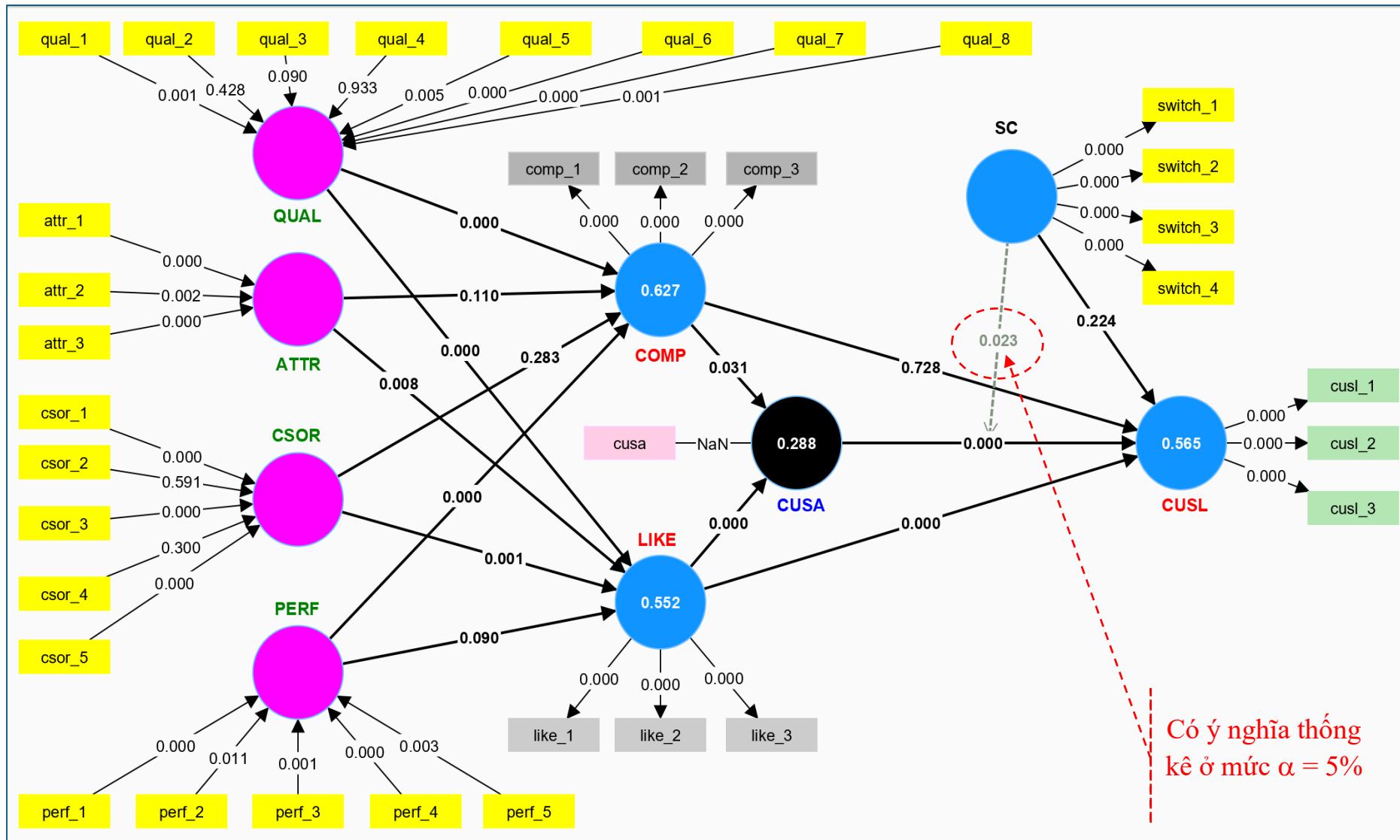
Lưu ý: Nên xem hướng dẫn tình huống ở [Case Study, Chapter 7](#) (theo SmartPLS 4), còn ở Chapter 7 trong ([Hair và cộng sự, 2022](#)) hướng dẫn theo SmartPLS 3.

Các bước thực hiện trên SmartPLS 4 như sau:

- Mở mô hình Extended model
- Chọn “**Select**” để tạo thêm một biến điều tiết (với SC thì nhấp đúp chuột tại vị trí thích hợp, sau đó chọn các biến chỉ báo như hướng dẫn ở các phần trước; với hai biến còn lại thì chỉ cần chọn – giữ chuột và kéo đến vị trí thích hợp).
- Chọn “**Connect**” → nhấp chuột vào biến điều tiết, giữ và kéo đến cấu trúc CUSL, sau đó thả ra; tiếp tục nhấp chuột vào biến điều tiết, giữ và kéo đến đường dẫn từ CUSA → CUSL, đặt ở bất kỳ vị trí nào, sau đó thả ra (làm tương tự cho các cấu trúc khác để mở rộng bài tập thực hành).
- Uớc lượng mô hình với thuật toán **PLS-SEM** (thực hiện nhanh các kiểm định đối với mô hình đo lường kết quả đối với biến SC – các biến servicetype và serviceprovider thì không cần kiểm định) → **Kiểm tra xem kết quả có giống với Exhibit 7.27** ([Hair và cộng sự, 2022: trang 264](#)) hoặc **Exhibit A7.4** ([Case Study, Chapter 7: trang 6](#))?
- (Kết quả PLS-SEM): **Final results** → **Simple slope analysis** để xem đồ thị về mối quan hệ giữa CUSA và CUSL khi biến điều tiết tăng/giảm một đơn vị độ lệch chuẩn.
- (Kết quả PLS-SEM): **Quality criteria** → **f-square** độ lớn ảnh hưởng của biến tương tác.
- Uớc lượng lại mô hình với thuật toán **Bootstrapping** để kiểm tra mức ý nghĩa của (các) hệ số tương tác, cũng như phân tích các loại tổng ảnh hưởng và tổng ảnh hưởng gián tiếp (nếu cần).
- Thử đi thử lại với các biến tương tác khác nhau trong mô hình.
- Thực hiện lại quy trình với biến điều tiết là loại dịch vụ và nhà cung cấp, và sau đó là kết hợp các biến điều tiết.

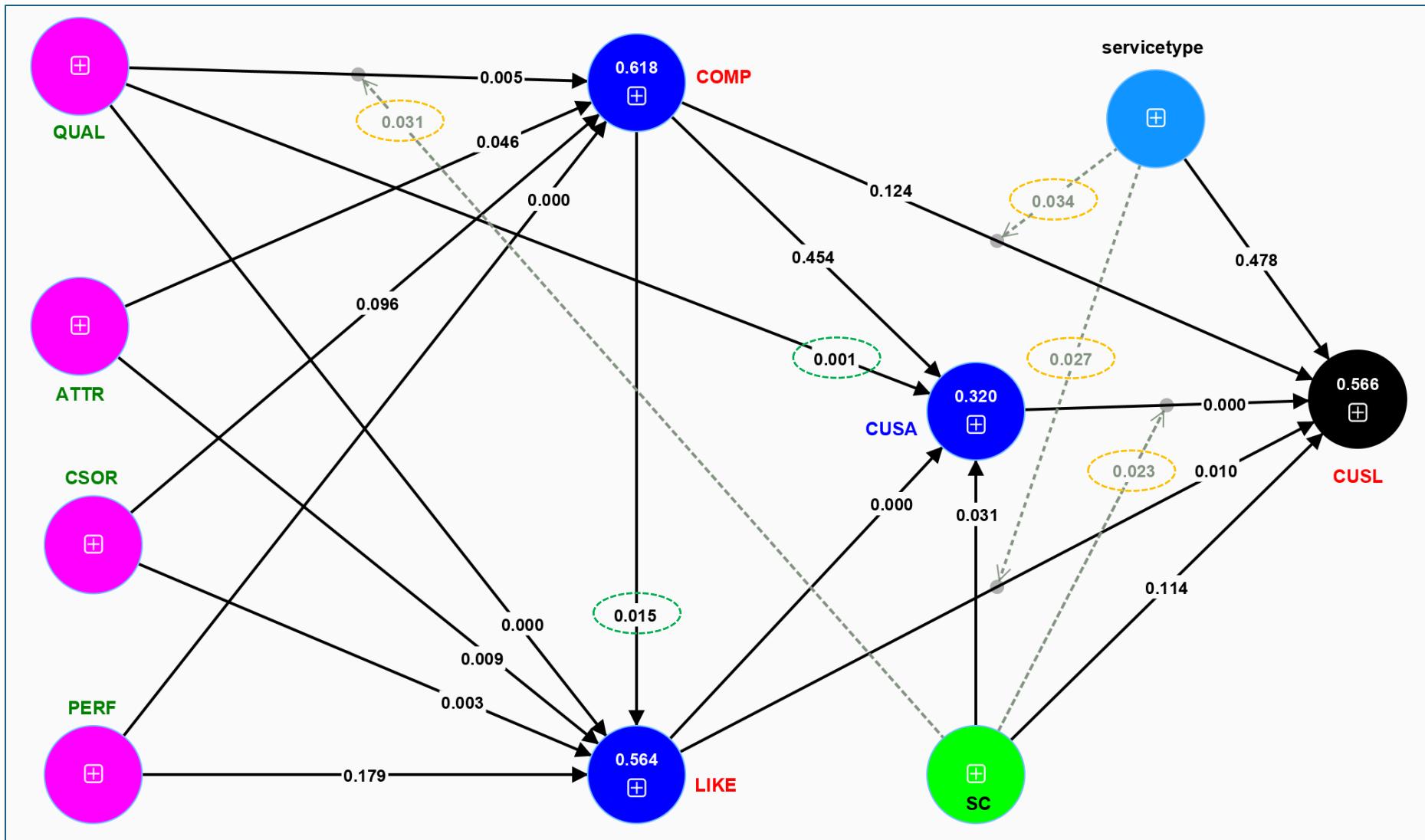


Hình 66. Kết quả ước lượng mô hình với biến điều tiết SC theo PLS-SEM.

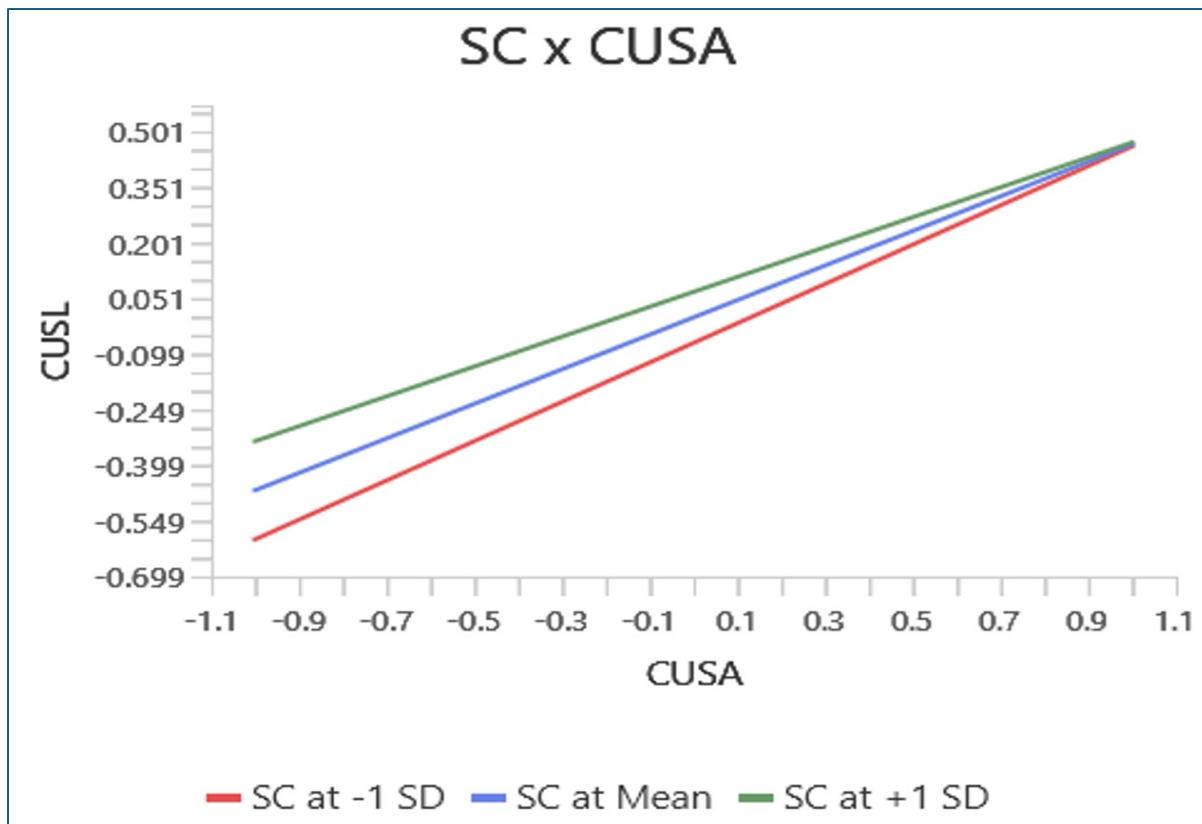


Có ý nghĩa thống kê ở mức $\alpha = 5\%$

Hình 67. Kết quả ước lượng mô hình với biến điều tiết SC theo Bootstrapping (**p-values**).



Hình 68. Kết quả ước lượng mô hình với biến điều tiết SC và Servicetype cuối cùng theo Bootstrapping (**p-values**).



Hình 69. Đồ thị tác động điều tiết của SC lên mối quan hệ CUSA và CUSL.



Hình 69. Đồ thị tác động điều tiết của Servicetype lên mối quan hệ LIKE và CUSL.

8. TỔNG KẾT

Chúng ta đã hoàn thành các nội dung cơ bản của mô hình PLS-SEM như cách ước lượng và kiểm định thống kê với phần mềm SmartPLS 4, cũng như cách giải thích kết quả của các mô hình đo lường và mô hình cấu trúc. **SmartPLS 4** giúp ta xây dựng một mô hình PLS-SEM một cách dễ dàng; ước lượng các mô hình PLS-SEM một cách nhanh chóng, nhất là kỹ thuật mô phỏng bằng bootstrapping; đánh giá các mô hình một cách thuận tiện; và trình bày các kết quả ước lượng một cách trực quan, sinh động.

Để có sự hiểu biết thấu đáo về mô hình PLS-SEM, ta cần khắc ghi các vấn đề sau đây. **1)** Hiểu thế nào là các biến chỉ báo hay biến quan sát, và các biến này được đo lường và thu thập như thế nào; **2)** Hiểu thế nào là các biến tiềm ẩn hay cấu trúc, tại sao sử dụng chúng, và chúng liên hệ như thế nào với các biến chỉ báo hay biến quan sát; **3)** Phân biệt giữa mô hình đo lường và mô hình cấu trúc; trong mô hình đo lường thì cần hiểu rõ sự khác biệt giữa mô hình đo lường kết quả và mô hình đo lường nguyên nhân cả về cách biểu diễn trên đồ thị cũng như dạng phương trình hồi quy; phân biệt giữa hệ số tải ngoài, trọng số ngoài, và hệ số đường dẫn; và cách giải thích hệ số đường dẫn với lựa chọn chuẩn hóa khi ước lượng mô hình; **4)** Biết được quy trình đánh giá một mô hình PLS-SEM từ mô hình đo lường kết quả đến mô hình đo lường nguyên nhân và sau cùng là mô hình cấu trúc, các nội dung quan trọng cần đánh giá trong mỗi mô hình là gì, so sánh lựa chọn mô hình cấu trúc phù hợp với dữ liệu sẵn có, cách thực hiện quy trình đánh giá các mô hình trên SmartPLS 4, các kết quả chính cần trình bày cho từng mô hình, khi nào nên dùng thuật toán PLS-SEM và Bootstrapping, và cần lưu ý đối với các giá trị bị thiếu trong tập dữ liệu; **5)** Trong mô hình cấu trúc cần phải phân biệt được các biến tiềm ẩn/cấu trúc ngoại sinh – nội sinh, độc lập – phụ thuộc, tiền vi/dự đoán – mục tiêu; thế nào là biến tiềm ẩn/cấu trúc trung gian, ảnh hưởng trực tiếp, (tổng) ảnh hưởng gián tiếp, tổng ảnh hưởng, và các thông kê kiểm định của các loại ảnh hưởng này là gì; **6)** Cuối cùng, phân biệt giữa ảnh hưởng trung gian và ảnh hưởng điều tiết kể cả cách biểu diễn trên đồ thị cũng như phương trình hồi quy; các loại biến điều tiết thường dùng là gì, tại sao sử dụng biến điều tiết; biến tương tác là gì và các cách tạo biến tương tác trong các mô hình PLS-SEM; cách tạo các mô hình điều tiết trên SmartPLS 4, ước lượng, kiểm định, và giải thích kết quả mô hình điều tiết.

Trong tài liệu hướng dẫn, chúng ta chỉ sử dụng tập dữ liệu “**Corporate reputation data**” để thực hành lại các tình huống được trình bày trong giáo trình của [Hair và cộng sự \(2022, 2018\)](#). Tuy nhiên, để nghiên cứu và thực hành theo hai giáo trình này ta phải tốn rất nhiều thời gian. Trong tài liệu hướng dẫn này, tôi trình bày lại các tình thực hành một cách xúc tích và trực quan hơn nên sẽ giúp ích cho bạn khá nhiều trong quá trình tự thực hành của mình. Hãy cài phần mềm, tải tập dữ liệu, và thực hiện từng bước theo hướng dẫn. Ở mỗi bước đánh giá mô hình, tôi tóm tắt các nội dung lý thuyết trọng tâm trước khi hướng dẫn từng thao tác thực hiện trên SmartPLS 4. Khi cần hiểu rõ hơn một vấn đề gì, bạn có thể tham khảo các giáo trình hoặc các tài liệu liên quan. Với giáo trình của [Hair và cộng sự \(2018\)](#) [bản dịch tiếng Việt, cũng sử dụng tập dữ liệu “Corporate reputation data”], các tác giả giới thiệu một số nội dung nâng cao; và tôi

nghĩ đây cũng là một nguồn tham khảo hữu ích. Một tài liệu khác của [Mehmetoglu và Venturini \(2021\)](#) cũng nên tham khảo và thực hành với một số tệp dữ liệu mới¹⁷. Tuy nhiên, giáo trình này hướng dẫn thực hành các mô hình PLS-SEM trên phần mềm Stata và R. Tuy nhiên, bạn sẽ dễ dàng thực hiện lại các bài tập này trên SmartPLS 4 và sau đó so sánh với các kết quả được trình bày trong tài liệu. Ngoài ra, trên trang chủ của SmartPLS¹⁸ có khá nhiều tài liệu tham khảo chất lượng và một số dự án (kèm dữ liệu) PLS-SEM thích hợp cho việc thực hành của bạn. Cuối cùng, bạn hãy đọc các bài nghiên cứu đăng trên các tạp chí danh tiếng, tìm “**Data in Brief**”, bạn sẽ có thể tải tệp dữ liệu (thường ở dạng **.csv**) và kèm phần mô tả các biến (thường ở dạng PDF). Cách này vừa giúp bạn hiểu sâu lý thuyết, nâng cao các kỹ năng học thuật, vừa có thể đối chiếu bài thực hành của bạn với các kết quả của bài nghiên cứu.

Tôi hy vọng tài liệu này hữu ích cho bạn trong việc tự học các mô hình PLS-SEM với phần mềm SmartPLS 4 khá phổ biến hiện nay. Tài liệu này tôi biên soạn dựa trên các giáo trình của [Hair và cộng sự \(2022, 2018\)](#) nhằm mục đích hỗ trợ sinh viên trong việc học tập các học phần liên quan đến phân tích dữ liệu và nghiên cứu, chứ không vì các mục đích thương mại. Trong một số phần, đoạn của tài liệu này, tôi tạm dịch từ các giáo trình gốc. Nếu như Việt hóa chưa chuẩn thì bạn có thể tìm và đọc các tài liệu gốc (có ghi chú số trang trong trích dẫn) để tránh hiểu sai vấn đề. Xin chân thành cảm ơn./.

¹⁷ <https://github.com/sergioventurini/SEMwPLS/>: Chọn Chapter → Data, tải về (nên chọn các tệp dữ liệu dạng **.csv** thay vì **.dta** để khỏi phải chuyển đổi dữ liệu).

¹⁸ <https://www.smartpls.com/documentation>: Chọn các Sample projects.

TÀI LIỆU ĐỌC THÊM

- Asteriou, D., & Hall, S. G. (2021). *Applied econometrics*. Bloomsbury Publishing.
- Becker, J.-M., Ringle, C. M., Sarstedt, M., & Völckner, F. (2015). How collinearity affects mixture regression results. *Marketing Letters*, 26(4), 643–659.
- Bollen, K. A., & Diamantopoulos, A. (2017). In defense of causal-formative indicators: A minority report. *Psychological Methods*, 22(3), 581–596.
- Chin, W. W. (1998). The partial least squares approach to structural equation modeling. In G. A. Marcoulides (Ed.), *Modern methods for business research* (pp. 295–358). Mahwah, NJ: Erlbaum.
- Chin, W. W. (2010). How to write up and report PLS analyses. In V. Esposito Vinzi, W. W. Chin, J. Henseler, & H. Wang (Eds.), *Handbook of partial least squares: Concepts, methods and applications in marketing and related fields* (Springer Handbooks of Computational Statistics Series, Vol. II, pp. 655–690). Berlin: Springer.
- Chin, W. W., Cheah, J.-H., Liu, Y., Ting, H., Lim, X.-J., & Cham, T. H. (2020). Demystifying the role of causal-predictive modeling using partial least squares structural equation modeling in information systems research. *Industrial Management & Data Systems*, 120(12), 2161–2209.
- Chin, W. W., Marcolin, B. L., & Newsted, P. R. (2003). A partial least squares latent variable modeling approach for measuring interaction effects: Results from a Monte Carlo simulation study and an electronic-mail emotion/adoption study. *Information Systems Research*, 14(2), 189–217.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Mahwah, NJ: Erlbaum.
- Danks, N., Ray, S., & Shmueli, G. (2017). Evaluating the predictive performance of composites in PLS path modeling. Working Paper, available at SSRN: <https://ssrn.com/abstract=3055222>.
- Franke, G., & Sarstedt, M. (2019). Heuristics versus statistics in discriminant validity testing: A comparison of four procedures. *Internet Research*, 29(3), 430–447.
- Gujarati, D. (2014). *Econometrics by example*. Bloomsbury Publishing.
- Gujarati, D. N., & Porter, D. C. (2009). *Basic econometrics*. McGraw-Hill.
- Hair, J. F., & Sarstedt, M. (2021). Explanation plus prediction – the logical focus of project management research. *Project Management Journal*.
- Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: Indeed a silver bullet. *Journal of Marketing Theory and Practice*, 19(2), 139–151.
- Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, 31(1), 2–24.

- Hair, J. F., Sarstedt, M., Ringle, C. M., & Gudergan, S. P. (2018). Advanced issues in partial least squares structural equation modeling (PLS-SEM). Thousand Oaks, CA: Sage.
- Hair, J., Hair Jr, J. F., Sarstedt, M., Ringle, C. M., & Gudergan, S. P. (2018). *Advanced issues in partial least squares structural equation modeling*. Sage publications (Bản dịch tiếng Việt).
- Hair, J., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2022). *A primer on partial least squares structural equation modeling (PLS-SEM)*. SAGE Publications, Incorporated.
- Hayes, A. F. (2015). An index and test of linear moderated mediation. *Multivariate Behavioral Research*, 50(1), 1–22.
- Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York, NY: Guilford.
- Helm, S., Eggert, A., & Garnefeld, I. (2010). Modelling the impact of corporate reputation on customer satisfaction and loyalty using PLS. In V. Esposito Vinzi, W. W. Chin, J. Henseler, & H. Wang (Eds.), *Handbook of partial least squares: Concepts, methods and applications in marketing and related fields* (Springer Handbooks of Computational Statistics Series, Vol. II, pp. 515–534). Berlin: Springer.
- Henseler, J., & Fassott, G. (2010). Testing moderating effects in PLS path models: An illustration of available procedures. In V. Esposito Vinzi, W. W. Chin, J. Henseler, & H. Wang (Eds.), *Handbook of partial least squares: Concepts, methods and applications in marketing and related fields* (Springer Handbooks of Computational Statistics Series, Vol. II, pp. 713–735). Berlin: Springer.
- Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, 43(1), 115–135.
- Hill, R. C., Griffiths, W. E., & Lim, G. C. (2018). *Principles of econometrics*. John Wiley & Sons.
- Homburg, C., & Giering, A. (2001). Personal characteristics as moderators of the relationship between customer satisfaction and loyalty—An empirical analysis. *Psychology & Marketing*, 18(1), 43–66.
- Malhotra, N. K. (2020). *Marketing research: An applied orientation*, Seventh Edition, Pearson.
- Mason, C. H., & Perreault, W. D. (1991). Collinearity, power, and interpretation of multiple regression analysis. *Journal of Marketing Research*, 28(3), 268–280.
- Mehmetoglu, M., & Venturini, S. (2021). *Structural equation modelling with partial least squares using Stata and R*. Crc Press.
- Mooi, E., Sarstedt, M., Mooi-Reci, I. (2018). *Market research: The process, data, and methods using Stata*. Springer.

- Muller, D., Judd, C. M., & Yzerbyt, V. Y. (2005). When moderation is mediated and mediation is moderated. *Journal of Personality and Social Psychology*, 89(6), 852–863.
- Nunan, D., Birks, D. F., & Malhotra, N. K. (2020). *Marketing research: Applied insight*. Pearson.
- Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers*, 36(4), 717–731.
- Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in simple and multiple mediator models. *Behavior Research Methods*, 40(3), 879–891.
- Preacher, K. J., Rucker, D. D., & Hayes, A. F. (2007). Assessing moderated mediation hypotheses: Theory, methods, and prescriptions. *Multivariate Behavioral Research*, 42(1), 185–227.
- Ramlall, I. (2016). *Applied structural equation modelling for researchers and practitioners: Using R and Stata for behavioural research*. Emerald Group Publishing Limited.
- Rigdon, E. E. (2012). Rethinking partial least squares path modeling: In praise of simple methods. *Long Range Planning*, 45(5–6), 341–358.
- Sarstedt, M., Ringle, C. M., & Hair, J. F. (2017). Partial least squares structural equation modeling. In C. Homburg, M. Klarmann, & A. Vomberg (Eds.), *Handbook of Market Research*. Cham: Springer.
- Sarstedt, M., Ringle, C. M., Henseler, J., & Hair, J. F. (2014). On the emancipation of PLS-SEM: A commentary on Rigdon (2012). *Long Range Planning*, 47(3), 154–160.
- Sarstedt, M., Ringle, C. M., Smith, D., Reams, R., & Hair, J. F. (2014). Partial least squares structural equation modeling (PLS-SEM): A useful tool for family business researchers. *Journal of Family Business Strategy*, 5(1), 105–115.
- Shmueli, G., Ray, S., Velasquez Estrada, J. M., & Chatla, S. B. (2016). The elephant in the room: Evaluating the predictive performance of PLS models. *Journal of Business Research*, 69(10), 4552–4564.
- Studenmund, A. H. (2017). *Using econometrics: A practical guide*. Pearson.

<https://www.pls-sem.net/downloads/>

<https://www.smartpls.com/downloads/>