UN SYNTHÈSE SUR LES ÉQUATIONS STRUCTURELLES : MÉTHODES D'ESTIMATION ET INDICES D'ADÉQUATION

PAR
STELLA GURRERI

© S. Gurreri, 2011

© Stella Gurreri, 14/02/2012

Les deux méthodes les plus utilisées pour tester un effet de médiation sont la régression multiple et l'équation structurelle (Li, 2011). Chacune de ces analyses possède des avantages et des inconvénients et s'applique davantage dans certains contextes.

De façon générale, le principal avantage de la régression multiple est qu'elle est plus facile à réaliser que l'équation structurelle (Li, 2011). Par contre, cette méthode comporte plusieurs inconvénients (Li, 2011). Avec les régressions multiples, il est très difficile de tester l'effet de plusieurs médiateurs sur la même variable dépendante. De plus, dans la régression, il est impossible de distinguer les résidus provenant des erreurs de mesure et ceux provenant d'autres causes. Cette méthode ne permet donc pas de contrôler les erreurs de mesures. Finalement, les variables utilisées dans la régression doivent être exclusivement de type "observées" puisqu'il est impossible d'utiliser des variables latentes (concepts représentés par plusieurs indicateurs ou mesures). En ce qui concerne la régression multiple, les postulats de base à respecter sont les suivants: distributions normales, résidus indépendants, absence de multicollinéarité, modèle correctement spécifié et absence d'erreur de mesure (Li, 2011).

De façon générale, l'équation structurelle serait la méthode la plus efficace pour tester un effet de médiation (Li, 2011). L'équation structurelle comprend un ensemble de techniques statistiques qui permet d'analyser un ensemble de relations entre une ou plusieurs variables indépendantes et une ou plusieurs variables dépendantes et permet d'analyser des relations complexes lorsque le phénomène est multidimensionnel (Tabachnick et Fidell, 2007). Cette procédure d’analyse permet de tester simultanément toutes les associations spécifiées dans un modèle incluant la force d'association entre les indicateurs observés et les concepts latents (modèle de mesure), le lien allant des variables indépendantes aux médiateurs et à la variable dépendante ainsi que le lien entre les médiateurs et la variable dépendante (modèle structurel) (Li, 2011). Cette méthode peut être utilisée pour tester un modèle, tester des hypothèses spécifiques à un modèle, modifier un modèle existant ou encore tester un ensemble de modèles reliés (Tabachnick et Fidell, 2007).

L'utilisation des équations structurelles a de multiples avantages (Li, 2011). Les avantages majeurs de cette méthode, comparativement aux régressions multiples, sont la possibilité de tester tous les effets médiateurs simultanément et le fait de prendre en compte l'erreur de mesure.

© Stella Gurreri, 14/02/2012

L'utilisation de plusieurs indicateurs aide à réduire l'impact des erreurs de mesure à l'aide de l'inclusion d'un modèle de mesure dans l'équation structurelle (Li, 2011). De plus, il est facile d'inclure plusieurs médiateurs dans les analyses et il y a aussi la possibilité de spécifier des relations réciproques dans le modèle. Le principal postulat à respecter pour l'équation structurelle est la normalité multivariée qui requiert que toutes les distributions univariées soient normales ainsi que toutes les distributions jointes de n'importe quelle paire de variables (Li, 2011; Tabachnick et Fidell, 2007; Kline, 2005). Il est à noter que les équations structurelles assument que les données manquantes peuvent être ignorées, c'est-à-dire que les données manquantes sont distribuées au hasard (Li, 2011). En fait, la plupart des logiciels faisant ce type d'analyse utilise la fonction "listwise" pour gérer les données manquantes, ce qui peut faire en sorte que les données ne soient pas généralisables à une population plus large si les données manquantes sont dans les faits non distribuées au hasard (Li, 2011). Les équations structurelles assument aussi qu'il y a absence de multicollinéarité et de mauvaise spécification du modèle.

Le postulat le plus important est la normalité multivariée. Puisqu'il est presque impossible de vérifier ce postulat, les auteurs (Tabachnick et Fidell, 2007; Kline, 2005) indiquent qu'il faut vérifier les valeurs aberrantes et la normalité des distributions.

Par contre, une méthode d'estimation tenant compte de la possibilité de la violation du postulat de normalité peut être utilisée: la méthode de Maximum de vraisemblance robuste (Robust Maximum Likelihood) puisqu'elle est la méthode conseillée dans la section Aide de Lisrel (version 8.8) lorsque les variables sont continues mais non-normales. De plus, des auteurs ont conclu que pour des modèles complexes, sous une condition de non-normalité, cette méthode donne relativement de bonnes propriétés statistiques comparativement aux autres méthodes d'estimation (Moindres carrés généralisés, Moindres carrés pondérés, etc.) (Boomsma et Hoogland, 2001). La méthode des Moindres carrés pondérés serait davantage recommandée pour des variables observées discrètes ou ordinales afin d'augmenter la fiabilité de l'estimation des paramètres (Li, 2011).

© Stella Gurreri, 14/02/2012

Selon Tabachnick et Fidell (2007), la première étape des analyses par équations structurelles est de vérifier si le modèle est identifié. Pour identifier le modèle, il faut examiner le ratio entre les « éléments d'information » – les variances et covariances entre les variables mesurées -- et les « paramètres à estimer », soit les divers coefficients du modèle. S'il y a plus d'éléments d'information que de paramètres à estimer le modèle est donc suridentifié, une condition nécessaire pour procéder à l'analyse.

Ensuite, il faut examiner le modèle de mesure. Il est nécessaire d'établir une échelle pour chaque variable latente et d'évaluer l'identifiabilité de la partie « mesure » du modèle. Pour ce faire, il est recommandé de fixer le coefficient de régression d'une des variables mesurées de chacune des variables latentes à 1. L'indicateur le plus fortement corrélé à la variable latente devrait être fixé à 1.

L'étape suivante consiste à examiner les résultats afin de s'assurer qu'il n'y a aucun avertissement et que le modèle est adéquat. En fait, une équation structurelle vérifie si et comment les relations prédites par un modèle théorique ne sont pas contredites par les données empiriques ; c'est ce que l'on appelle l'adéquation entre le modèle théorique et les données. À partir des résultats, il faut évaluer l'adéquation du modèle et au besoin modifier les liens (en ajouter ou en supprimer) et refaire l'analyse jusqu'à ce que le modèle soit considéré satisfaisant. En effet, la vérification de l'adéquation du modèle est la première étape dans l'analyse des résultats des équations structurelles (Tabachnick et Fidell, 2007; Kline, 2005).

Pour évaluer l'adéquation, il faut examiner différents indices. Plusieurs indices existent et Lisrel en fournit plusieurs dans ses résultats. Les écrits sur le sujet aident à mieux s'y retrouver.

Parmi les nombreux indices d'adéquation de modèle, le chi-carré est un des premiers qui a été utilisé (Sharma, Mukherjee, Kumar et Dillon, 2005). Cet indice, comme plusieurs autres, est sensible à la taille de l'échantillon et, pour une même taille d'échantillon, le nombre d'indicateurs influence aussi l'indice. Dans le cadre des équations structurelles, le chi-carré ne doit pas être significatif, car cela confirme une différence significative entre le modèle et les données. Une grande taille d'échantillon fait en sorte qu'une différence minime entre le modèle et les données

© Stella Gurreri, 14/02/2012

donne un chi-carré significatif. Il faut donc se fier à d'autres indices d'adéquation du modèle (Fan, Thompson et Wang, 1999). Il n'y aurait pas d'indice idéal et c'est pour cela que plusieurs indices doivent être analysés (Fan, Thompson et Wang, 1999).

Un premier type d'indice est dérivé de la reproduction de la matrice de covariance qui essaie d'évaluer le degré avec lequel la matrice reproduite basée sur le modèle ressemble à la matrice de covariance de l'échantillon. Le Goodness of Fit et le Adjusted Goodness of Fit en sont des exemples (Fan, Thompson et Wang, 1999). Un deuxième type sont les indices de comparaison qui évaluent l'adéquation en comparant un modèle donné avec un modèle nul plus restreint (le modèle nul assume qu'il n'y a aucune relation entre les indicateurs du modèle). Des exemples sont le Normed Fit Index et le Nonnormed Fit Index qui tiennent compte des degrés de liberté (Fan, Thompson et Wang, 1999). Le troisième type d'indices évalue la parcimonie. Ces indices seront moins élevés si le modèle est trop élaboré; ils prennent en compte les degrés de liberté. Il y a aussi les indices d'adéquation qui utilisent les statistiques de noncentralité de la distribution du chi-deux. Dans cette famille, on retrouve le Comparative Fit Index qui évalue l'adéquation du modèle comparativement à un modèle de base (Fan, Thompson et Wang, 1999).

D'autres indices sont souvent retrouvés dans les écrits. Le Root Mean Square Error of Approximation (RMSEA) qui estime le manque d'adéquation d'un modèle comparativement à un modèle parfait (Tabachnick et Fidell, 2007). Les indices portant sur les résidus tels que le Root Mean Square Residual (RMR) et le Standardized Root Mean Square Residual (SRMR). Ces derniers comparent la différence moyenne entre la variance et la covariance de l'échantillon et celle de la population estimée (Tabachnick et Fidell, 2007).

Certains indices seraient plus performants que d'autres. Un bon indice ne devrait pas varier excessivement selon la taille de l'échantillon ni selon la méthode d'estimation, mais dans les faits peu d'indices répondent complètement à ces critères (Fan, Thompson et Wang, 1999).

Une étude très citée dans les écrits scientifiques, celle de Hu et Bentler (1998), recommande de façon générale de ne pas utiliser les indices suivants: le Normed Fit Index (NFI), le Goodness of Fit Index (GFI) et le Adjusted Goodness of Fit Index (AGFI) qui est ajusté selon le nombre de

© Stella Gurreri, 14/02/2012

paramètres dans le modèle. Ces indices ne sont pas recommandés pour plusieurs raisons: Ils ne sont pas sensibles à la mauvaise spécification du modèle et ils sont sensibles à la taille de l'échantillon avec certaines méthodes d'estimation. De façon générale, ces auteurs recommandent plutôt l'utilisation du Comparative Fit Index (CFI) en raison de sa sensibilité à une mauvaise spécification du modèle ainsi qu'à sa moindre sensibilité à la méthode d'estimation, à la distribution des données et à la taille de l'échantillon. Les auteurs recommandent aussi le Root Mean Square Error of Approximation (RMSEA) et le Standardized Root Mean Square Residual (SRMR). Plus spécifiquement, lorsqu'une méthode d'estimation de maximum de vraisemblance est utilisée, les auteurs recommandent d'utiliser au moins deux indices d'adéquation, soit le SRMR en combinaison avec le CFI ou le RMSEA (parmi ceux disponibles avec le logiciel Lisrel).

Ces recommandations ressemblent à celles données par Sharma, Mukherjee, Kumar et Dillon (2005). Selon les résultats de leur étude, le Tucker Lewis Index (TLI) et le Relative Noncentrality Index (RNI) sont les indices qui performent le mieux suivis par le Normed Non Centrality Parameter (NNCP) et le RMSEA. Quant au GFI, il a montré de moins bons résultats. C'est l'indice qui a le moins bien performé puisqu'il était affecté grandement par la taille de l'échantillon et par le nombre d'indicateurs et qu'il n'était pas vraiment sensible à une mauvaise spécification du modèle. Les auteurs suggèrent donc de ne pas utiliser cet indice pour évaluer l'adéquation du modèle. Quant au RMSEA, il serait un meilleur indice que le GFI, mais ne serait pas aussi bon que d'autres indices. Les auteurs recommandent son utilisation en combinaison avec d'autres indices.

Ces résultats ne font pas l'unanimité puisque d'autres auteurs croient que le GFI et le AGFI seraient tout de même plus sensibles à une mauvaise spécification du modèle que des indices tels que CFI, NNFI et NFI (Fan, Thompson et Wang, 1999). Effectivement, selon ces auteurs, le RMSEA, suivi du GFI et du AGFI seraient sensibles à une mauvaise spécification du modèle et seraient moins influencés par la méthode d'estimation. Par contre, les auteurs ne recommandent pas plus l'utilisation du GFI ou de l'AGFI. Le RMSEA étant peu influencé par la taille de l'échantillon, le RMSEA en combinaison avec le CFI et le NNFI seraient les meilleurs choix

© Stella Gurreri, 14/02/2012

d'indices d’adéquation. Le GFI et le AGFI auraient moins bien performé en raison de leur variation selon la taille de l'échantillon.

D'après les recommandations précédentes et les indices disponibles dans Lisrel, le chi-carré, le RMSEA, le CFI, NNFI et le SRMR semblent les meilleurs indices à utiliser afin d'estimer l’adéquation d’un modèle. Les auteurs indiquent que pour un indice d'adéquation se situant entre 0 et 1, une valeur au-dessus de .90 est acceptable, une valeur sous la barre du 0,90 étant considérée inacceptable (Sharma, Mukherjee, Kumar et Dillon, 2005; Hu et Bentler, 1998). Une valeur près de 0.95 indiquerait une bonne adéquation (Tabachnick et Fidell, 2007). Plus spécifiquement, un CFI en haut de 0.95 couplé à un RMSEA en bas de 0,06 et à un SRMR en bas de 0.08 sont considérés comme des indicateurs d'une bonne adéquation entre le modèle postulé et les données observées (Sharma, Mukherjee, Kumar et Dillon, 2005; Fan, Thompson et Wang, 1999; Hu et Bentler, 1998).

© Stella Gurreri, 14/02/2012

Références

BoomSma, A., et Hoogland, J.J. The Robustness of LISREL modeling revisited In Cudeck, R., Du Toit, S., Sorbom, D.Ed Structural Equation Modeling: Present and Future. Scientific Software International (2001).

Fan, X., Thompson, B., et Wang, L. (1999). Effects of sample size, estimation methods, and model specification on structural equation modeling fit indexes, *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 56-83.

Hu, L., Bentler, P. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification, *Psychological Methods,* 3(4), 424-453.

Kline, R. (2005). Principles and Practice of Structural Equation Modeling, 2nd ed. New York: The Guilford Press.

Li, D.S. (2011). Testing mediation using multiple regression and structural equation modeling analyses in secondary data, *Evaluation Review*, 35(3), 240-268.

Sharma, S., Mukherjee, S., Kumar, A., et Dillon, W.R. (2005). A simulation study to investigate the use of cutoff values for assessing model fit in covariance structure models, *Journal of Business Research*, 58, 935-943.

Tabachnick, B. G., et Fidell, L. S. (2007). Using Multivariate Statistics , 5th ed. Boston: Allyn and Bacon.

© Stella Gurreri, 14/02/2012