MATHÉMATIQUES ET SCIENCES HUMAINES

CHRISTIAN HEUCHENNE

Effet des erreurs de mesure sur l'évaluation de l'efficacité d'un traitement

Mathématiques et sciences humaines, tome 122 (1993), p. 5-19

<http://www.numdam.org/item?id=MSH_1993__122__5_0>

Math. Inf. Sci. hum. (31e année, n° 122, *1993*, p. 5-19)

EFFET DES ERREURS DE MESURE SUR L'ÉVALUATION DE L'EFFICACITÉ D'UN TRAITEMENT

Christian HEUCHENNE[1]

RÉSUMÉ — *L'effet d'un traitement T sur une compétence Y peut être exprimé par le coefficient de régression partiel* β_{YT} *avec contrôle de la compétence initiale X. Quand Y et X sont mesurées avec erreurs par y et* $\vec{x}$, *cet effet se manifeste par le coefficient* β_{yT} *dans la régression de y sur T et* $\vec{x}$. *Le biais entre* β_{YT} *et* β_{yT} *est explicité, discuté et montré systématique si X et T sont corrélés. L'importance de bien spécifier le modèle, dont une condition nécessaire et suffisante d'identification est donnée, est mise en évidence.*

SUMMARY — Effect of measurement errors on the efficiency assessment of a treatment. *The effect of a treatment T on a skill can be conveyed by the partial regression coefficient* β_{YT} *with controlling the initial skill X. Where Y and X are measured with errors by y and* $\vec{x}$, *this effect is revealed by the coefficient* β_{yT} *of the regression of y onto T and* $\vec{x}$. *Bias between* β_{YT} *and* β_{yT} *is made explicit, discussed and shown one-sided if X and T correlate. Importance of good specification of the model, of which a neccessary and sufficient identification condition is given, is emphasized.*

1. EXPOSÉ DU PROBLÈME

Dans la littérature anglo-saxonne, les mises en garde contre l'usage irréfléchi de la technique de régression multiple pour évaluer l'efficacité d'un traitement pédagogique sont nombreuses [Cook, Campbell].

Dans un rapport fondé sur cette méthode [Mingat], il est avancé que la rééducation en GAPP (groupes d'aide psychopédagogique) a un effet nocif sur le niveau scolaire. Cette conclusion a été vivement contestée pour des raisons plus sentimentales ou sociales que scientifiques [Le Monde de l'Education].

L'article présent se veut une mise au point formelle de la question. Il a pour objectifs de démontrer que

a) pour juger correctement l'impact d'un traitement sur une compétence imprécisément mesurée, il faut soumettre les observations à un modèle qui prend compétences initiale et finale au titre de variables latentes ;
b) l'effet apparent biaise l'effet réel ainsi dégagé dans un sens privilégié quand la compétence initiale est entachée d'erreurs dans sa mesure et oriente la sélection des sujets traités, ce qui est

[1] Faculté de Psychologie et des Sciences de l'Éducation de l'Université de Liège.

toujours le cas si le traitement est une remédiation destinée aux individus les plus faibles ou un enrichissement destiné aux individus les plus forts.

Concrètement nous mettrons en question, après bien d'autres, l'usage répandu d'évaluer l'efficacité d'un traitement spécial T sur le rendement scolaire par la régression d'un posttest y sur T et divers prétests x_i, la conclusion se limitant à peu de choses près à interpréter le coefficient brut β_{yT} de y sur T avec contrôle des x_i. Le danger est trop grand en effet que les inévitables erreurs de mesure affectant y et les x_i apportent une distorsion importante et systématique de β_{yT} par rapport à un coefficient net reflétant le gain ou la perte dû à T de la compétence réelle, après neutralisation de la compétence initiale, c'est-à-dire le facteur mesuré par les x_i.

Nous remercions A. Grisay, chercheur en pédagogie expérimentale à l'Université de Liège, de nous avoir remis sur une piste de cet éternel problème en sciences humaines.

2. PRÉSENTATION DU MODÈLE

La question est : l'intervention d'un traitement extrascolaire (remédiation à l'intention des faibles, enrichissement à l'intention des forts) modifie-t-elle telle compétence des sujets ?

Soient T le traitement, X la compétence avant l'application de T, Y la compétence après cette application. La question pourrait être résolue sans problème en effectuant une régression de Y sur T et X

$$Y = \beta_{YT} T + \beta_{YX} X + \zeta \qquad (1)$$

C'est le coefficient β_{YT} qui est au centre de l'intérêt car il évalue l'influence spécifique de T sur Y, la compétence initiale X étant neutralisée. En effet, quand on retire de la compétence finale Y la part évidemment due à la compétence initiale (terme $\beta_{YX} X$) et l'impact d'autres variables déterminantes qui sont négligées, inconnues ou aléatoires (résidu ζ), il reste l'effet propre de T (terme $\beta_{YT}T$).

Selon que β_{YT} est positif, nul ou négatif, on en conclura que le traitement T est favorable, indifférent ou défavorable à la compétence.

En fait - et ceci nuance décidément la netteté de l'équation (1) - une compétence se révèle sous de multiples aspects et est donc mesurée par divers indicateurs plus ou moins faillibles. Dans le contexte présent, X (resp. Y) se manifeste sous forme de prétests (resp. posttests) x_i (resp. y_j) et la théorie classique des mesures congénériques [Bollen, p. 208] pose les relations

$$x_i = \lambda_i X + \delta_i \qquad (2)$$

$$y_j = \lambda_j Y + \varepsilon_j \qquad (3)$$

Les δ_i et ε_j sont les erreurs de mesure de la psychométrie. A un niveau plus fin, rien n'empêche même que les x_i ou y_j soient les items d'un test.

Schématiquement, ce que trop d'auteurs se bornent à faire, c'est la régression d'une mesure globale de Y

$$y = \sum_j p_j y_j \qquad (4)$$

sur T et les observations x_i

$$y = \beta_{yT} T + \sum_i \beta_{yi} x_i + \zeta' \qquad (5)$$

Ils interprètent alors le coefficient β_{yT} en ignorant l'existence des erreurs de mesure.

Ce n'est pas β_{yT} qu'il faut regarder, mais bien β_{YT} obtenu après avoir dégagé X et Y de la gangue de leurs mesures faillibles x_i et y_j. Autrement dit, X et Y doivent être considérées comme des variables latentes (facteurs). L'étude présente n'a évidemment de sens que si l'ensemble des x_i (resp. y_j) ramasse bien l'entièreté de ce qu'on met dans le concept de compétence initiale (resp. finale) ; la réalité de la variable construite X (resp. Y) en dépend.

Par contre, on peut croire que T n'est pas affectée d'erreur car cette variable résulte d'une imposition humaine volontaire. T peut être, par exemple, le temps d'administration du traitement. Une version plus rudimentaire, mais encore correcte, code la dichotomie traité ou non par deux valeurs arbitraires (usuellement 1 et 0).

On est amené naturellement à combiner trois modèles :
- une analyse factorielle des x_i qui en induit leur substance première, la compétence initiale X ;
- une analyse factorielle des y_j qui en induit Y;
- une régression de Y sur T et X.

On entre alors dans un modèle, dit de structure de covariances, qui permet d'estimer notamment le coefficient net β_{YT}. Les seules conditions essentielles imposées à ce modèle sont routinières :

1) non corrélation des erreurs de mesure aux variables fondamentales

$$\mathrm{cov}(\delta_i,Y) = \mathrm{cov}(\delta_i,T) = \mathrm{cov}(\delta_i,X) = 0 \qquad \text{pour tout } i$$
$$\mathrm{cov}(\varepsilon_j,Y) = \mathrm{cov}(\varepsilon_j,T) = \mathrm{cov}(\varepsilon_j,X) = 0 \qquad \text{pour tout } j$$

2) non corrélation du résidu aux explicateurs de Y dans (1)

$$\mathrm{cov}(\zeta,T) = \mathrm{cov}(\zeta,X) = 0$$

Cette dernière exigence n'est pas innocente. Elle postule en somme que tout déterminant de Y autre que T et X, placé par définition dans ζ, n'est associé ni à X, ni à T. C'est une simplification de la réalité dont il faut avoir conscience pour apprécier la démonstration qui suit. Le paragraphe 8 pallie cette faiblesse mais est moins décisif.

Notons que sont permises *a priori* des corrélations non nulles entre toutes les erreurs de mesure, y compris entre les δ_i des prétests et les ε_j des posttests. Le réalisme y invite car des traits différents, ici X et Y, peuvent être mesurés par une méthode identique ou similaire.

La métrique des variables Y, T, X (parce qu'elles sont latentes) et y (parce que les poids p_j de (4) ne sont définis qu'à une constante de proportionnalité près) doit être fixée ; elle le sera tout bonnement en demandant que ces variables soient réduites, c'est-à-dire de moyenne nulle et de variance unité. Une conséquence agréable de cette définition d'échelles est la possibilité de comparer directement β_{yT} et β_{YT}.

Enfin, pour simplifier les écritures et sans nuire à la généralité du raisonnement, toutes les variables observées x_i et y_j seront aussi supposées réduites. Toutes les covariances, à l'exception de celles portant sur une erreur de mesure, seront donc des corrélations.

De ces contraintes résulte déjà

$$1 = \lambda_y^2 + \mathrm{var}(\varepsilon)$$

en posant

$$\lambda_y = \sum_j p_j \lambda_j \qquad \varepsilon = \sum_j p_j \varepsilon_j$$

puisque, comme ses composants ε_j, ε est non corrélé avec Y de variance unité et que, par (4) et (3)

$$y = \left(\sum_j p_j \lambda_j\right) Y + \sum_j p_j \varepsilon_j = \lambda_y Y + \varepsilon$$

Dans la théorie classique de la mesure, λ_y^2 est la fidélité de y vis-à-vis de la compétence finale vraie Y.

3. RELATION ENTRE COEFFICIENTS BRUT ET NET

Jusqu'ici on a mis en lumière que l'influence de T et X sur Y passe, dégradée par l'imprécision des mesures, dans la régression de y sur T et l'ensemble des x_i (équation 5). Quelle est la relation entre le coefficient brut β_{yT} de (5) et le coefficient net β_{YT} de (1), et quels ingrédients entrent dans cette relation ?

La formule bien connue fournit

$$\beta_{yT} = \frac{\rho_{yT} - \vec{\rho}\,'_{xT} \Sigma_{xx}^{-1} \vec{\rho}_{xy}}{1 - \vec{\rho}\,'_{xT} \Sigma_{xx}^{-1} \vec{\rho}_{xT}} \tag{6}$$

où ρ_{yT} est la corrélation de y et T,

$\vec{\rho}_{xT}$ est le vecteur colonne des corrélations ρ_{iT} entre les x_i et T,

$\vec{\rho}\,'_{xT}$ est le vecteur ligne transposé de $\vec{\rho}_{xT}$,

Σ_{xx} est la matrice des corrélations ρ_{ik} de x_i et x_k,

$\vec{\rho}_{xy}$ est le vecteur colonne des corrélations ρ_{iy} entre les x_i et y.

En notant $\vec{\lambda}$ le vecteur des saturations λ_i de (2), les égalités

$$\rho_{iT} = \text{cov}(\lambda_i X + \delta_i, T) = \lambda_i \text{cov}(X,T) + \text{cov}(\delta_i,T) = \lambda_i \rho_{XT} \tag{7}$$

se résument en

$$\vec{\rho}_{xT} = \rho_{XT} \vec{\lambda}$$

où, cela va de soi, ρ_{XT} est la corrélation de X et T.
Pour le dénominateur de (6) on a donc

$$\vec{\rho}\,'_{xT} \Sigma_{xx}^{-1} \vec{\rho}_{xT} = \rho_{XT}^2 \vec{\lambda}\,' \Sigma_{xx}^{-1} \vec{\lambda}$$

La quantité $\vec{\lambda}\,' \Sigma_{xx}^{-1} \vec{\lambda}$ mérite d'être baptisée, — appelons la λ_x^2 —, car c'est la fidélité de l'ensemble des x_i par rapport à la variable latente X que ces indicateurs sont censés mesurer. En effet

$$\text{cov}(x_i X) = \text{cov}(\lambda_i X + \delta_i, X) = \lambda_i \text{cov}(X,X) + \text{cov}(\delta_i X) = \lambda_i$$

d'où l'on tire que le coefficient de détermination de X par les x_i est
(variance de X prise en compte par les x_i)/(variance de X) =

$$\overrightarrow{\text{cov}}\,'(x_i,X)\ \Sigma_{xx}^{-1}\ \overrightarrow{\text{cov}}\,(x_i,X)\,/1 = \vec{\lambda}\,' \Sigma_{xx}^{-1} \vec{\lambda}$$

Poursuivons l'exploration de (6).

$$\begin{aligned}\rho_{iy} &= \mathrm{cov}(x_i,y) = \mathrm{cov}(\lambda_i X + \delta_i, \lambda_y Y + \varepsilon)\\ &= \lambda_i \lambda_y \mathrm{cov}(X,Y) + \lambda_i \mathrm{cov}(X,\varepsilon) + \lambda_y \mathrm{cov}(\delta_i,Y) + \mathrm{cov}(\delta_i,\varepsilon)\\ &= \lambda_i \lambda_y \rho_{XY} + \mathrm{cov}(\delta_i,\varepsilon)\end{aligned}$$

en notant ρ_{XY} la corrélation de X et Y.

Ecrivant encore $\vec{\delta}$ le vecteur des erreurs δ_i de (2), la version vectorielle des équations précédentes est

$$\vec{\rho}_{xy} = \lambda_y \rho_{XY} \vec{\lambda} + \mathrm{cov}(\vec{\delta},\varepsilon)$$

L'expression $\vec{\rho}\,'_{xT} \Sigma^{-1}_{xx} \vec{\rho}_{xy}$ au numérateur de (6) est donc

$$\rho_{xT} \vec{\lambda}' \Sigma^{-1}_{xx} (\lambda_y \rho_{XY} \vec{\lambda} + \mathrm{cov}(\vec{\delta},\varepsilon)) = \rho_{XT} \lambda^2_x \lambda_y \rho_{XY} + \rho_{XT} \vec{\lambda}' \Sigma^{-1}_{xx} \mathrm{cov}(\vec{\delta},\varepsilon)$$

Reste à expliciter ρ_{yT} de (6), puis la méchante expression

$$\vec{\lambda}' \Sigma^{-1}_{xx} \mathrm{cov}(\vec{\delta},\varepsilon) .$$

Pour ρ_{yT} il vient immédiatement

$$\rho_{yT} = \mathrm{cov}(\lambda_y Y + \varepsilon, T) = \lambda_y \mathrm{cov}(Y,T) + \mathrm{cov}(\varepsilon,T) = \lambda_y \rho_{YT}$$

où, cela va de soi, ρ_{YT} est la corrélation de Y et T.

La dernière démarche est la plus subtile. Constatant que la partie "nette" $\vec{\lambda}X$ du vecteur $\vec{x}$ assemblant les composantes de (2) a été implicitement résumée en

$$\lambda_x X = \lambda^2_x X/\lambda_x = \vec{\lambda}' \Sigma^{-1}_{xx} (\vec{\lambda}X) / \lambda_x$$

on définira naturellement une erreur scalaire globale δ de l'ensemble des indicateurs x_i par

$$\delta = \vec{\lambda}' \Sigma^{-1}_{xx} \vec{\delta} / \lambda_x \qquad (8)$$

Les plus futés verront que ce qui est évoqué ainsi est la combinaison linéaire des x_i

$$\vec{\lambda}' \Sigma^{-1}_{xx} \vec{x} / \lambda_x = \vec{\lambda}' \Sigma^{-1}_{xx} (\vec{\lambda}X + \vec{\delta}) / \lambda_x = \lambda_x X + \delta$$

qui maximise la fidélité (ici λ^2_x) vis-à-vis de X. En d'autres mots, du point de vue régressif, le vecteur $\vec{x}$ des x_i est équivalent à cette unique variable réduite. D'ailleurs les incroyants seront convaincus par la réussite technique qui suit.

$$\rho_{ik} = \mathrm{cov}(x_i,x_k) = \mathrm{cov}(\lambda_i X + \delta_i, \lambda_k X + \delta_k) = \lambda_i\lambda_k + \mathrm{cov}(\delta_i,\delta_k) \quad (9)$$

soit, en notant $\theta_{\delta\delta}$ la matrice des covariances entre δ_i, On en déduit successivement

$$\Sigma_{xx} = \vec{\lambda}\,\vec{\lambda}' + \theta_{\delta\delta}$$

$$\vec{\lambda}'\,\Sigma_{xx}^{-1}\,\theta_{\delta\delta} = \vec{\lambda}'\,\Sigma_{xx}^{-1}\,(\Sigma_{xx} - \vec{\lambda}\,\vec{\lambda}') = \vec{\lambda}' - (\vec{\lambda}'\,\Sigma_{xx}^{-1}\vec{\lambda})\,\vec{\lambda}' = (1-\lambda_x^2)\,\vec{\lambda}'$$

et la variance de (8) :

$$\vec{\lambda}'\,\Sigma_{xx}^{-1}\,\theta_{\delta\delta} = \vec{\lambda}'\,\Sigma_{xx}^{-1}\,(\Sigma_{xx} - \vec{\lambda}\,\vec{\lambda}') = \vec{\lambda}' - (\vec{\lambda}'\,\Sigma_{xx}^{-1}\vec{\lambda})\,\vec{\lambda}' = (1-\lambda_x^2)\,\vec{\lambda}' \tag{10}$$

En désignant par $\rho_{\delta\varepsilon}$ la corrélation des erreurs scalaires δ et ε,

$$\vec{\lambda}'\,\Sigma_{xx}^{-1}\,\mathrm{cov}(\vec{\delta},\varepsilon) = \lambda_x\,\mathrm{cov}(\vec{\lambda}'\,\Sigma_{xx}^{-1}\vec{\delta}\,/\,\lambda_x,\,\varepsilon) = \lambda_x\,\mathrm{cov}(\delta,\varepsilon) = \lambda_x\sqrt{1-\lambda_x^2}\,\sqrt{1-\lambda_y^2}\,\rho_{\delta\varepsilon}$$

Portant tout cela dans (6) et tenant compte des expressions de ρ_{YT}, ρ_{XY} livrées par les équations, dites normales, résultant de covariances prises en (1) :

$$\rho_{YT} = \beta_{YT} + \beta_{YX}\,\rho_{XT} \qquad \rho_{XY} = \beta_{YT}\,\rho_{XT} + \beta_{YX} \tag{11}$$

il vient finalement

$$\beta_{yT} = \lambda_y\,\beta_{YT} + \frac{\rho_{XT}\sqrt{1-\lambda_x^2}\,\left(\lambda_y\sqrt{1-\lambda_x^2}\,\beta_{YX} - \lambda_x\sqrt{1-\lambda_y^2}\,\rho_{\delta\varepsilon}\right)}{1-\lambda_x^2\,\rho_{XT}^2}\,. \tag{12}$$

4. DISCUSSION DU BIAIS ENTRE LES DEUX COEFFICIENTS

Récrivons (12) sous la forme compacte

$$\beta_{yT} = \lambda_y\,\beta_{YT} + \frac{\rho_{XT}\,\tau}{1-\lambda_x^2\rho_{XT}^2} \tag{13}$$

où

$$\tau = \sqrt{1-\lambda_x^2}\,\left(\lambda_y\sqrt{1-\lambda_x^2}\,\beta_{YX} - \lambda_x\sqrt{1-\lambda_y^2}\,\rho_{\delta\varepsilon}\right)$$

dépend des paramètres de mesure λ_x, λ_y, $\rho_{\delta\varepsilon}$ et du taux linéaire β_{YX} de croissance de Y en X.

Dans cette formulation de β_{yT}, le terme $\lambda_y\beta_{YT}$ est clairement interprétable : il manifeste une atténuation de l'effet net β_{YT}, d'autant plus grande que y mesure mal la compétence finale (λ_y^2 est la fidélité de y par rapport à Y).

Plus énigmatique en (13) est le biais additif

$$\alpha = \frac{\rho_{XT}\,\tau}{1-\lambda_x^2\,\rho_{XT}^2}$$

dont le signe est celui de $\rho_{XT}\tau$ puisque $|\,\lambda_x\,| \leq 1$ (par (10)) et $|\,\rho_{XT}| < 1$.

Ce biais α est dû à la conjonction d'une corrélation ρ_{XT} entre compétence initiale et traitement, avec des erreurs dans la mesure de X (τ est généralement non nul quand la fidélité λ_x^2 est inférieure à 1).

Les circonstances où α s'évanouit sont intéressantes à scruter car (13) se réduit alors à

$$\beta_{yT} = \lambda_y \beta_{YT} \qquad \text{avec} \qquad 0 \leq \lambda_y \leq 1$$

ce qui entraîne que coefficients brut et net ont même signe et que $|\beta_{yT}| \leq |\beta_{YT}|$. Ceci permettrait de conclure sans ambiguïté à un effet nocif, nul ou bénéfique de T sur Y dont la magnitude serait cependant atténuée par l'infidélité de y.

Interprétons les deux cas où cela arrive.

1) $\rho_{XT} = 0$. Compétence de départ et traitement ne seraient pas corrélés. Ce phénomène ne s'observe qu'exceptionnellement dans un relevé passif de données naturelles.
Dans un contexte de plan expérimental, cette situation est cependant artificiellement créée par assignation au hasard des modalités du traitement aux individus, ce qui rend T indépendante de la variable X contemporaine. Ce constat est rassurant car il était escompté.

2) $\lambda_x^2 = 1$ qui annule τ et α. $\text{var}(\delta) = 1 - \lambda_x^2 = 0$ (équation 10) implique $\delta = \vec{\lambda}' \Sigma_{xx}^{-1} \vec{\delta} / \lambda_x = 0$ (équation 8), donc $\vec{\delta} = \vec{0}$.

La mesure de X serait réalisée fidèlement par ses indicateurs x_i, avec des erreurs δ_i nulles. Faut-il ajouter que c'est un idéal ?

Quand $\rho_{XT} \neq 0$ et $\lambda_x^2 < 1$, que peut-on dire du signe du biais additif α ? Ici se place le point crucial de la thèse que nous défendons : τ sera le plus souvent positif.

τ est négatif, nul ou positif selon que β_{YX} est inférieur, égal ou supérieur à $\rho_{\delta\varepsilon} \sqrt{\dfrac{\lambda_y^{-2} - 1}{\lambda_x^{-2} - 1}}$

Or dans les faits,

1) β_{YX} est certainement positif et vraisemblablement proche de 1 car, *à T constant*, à tout le moins la compétence finale Y est fonction croissante du niveau de départ X.

2) $\rho_{\delta\varepsilon}$ est borné supérieurement par 1 et souvent voisin de 0, car il n'y a guère de raisons que les erreurs de mesure aux prétests δ_i et aux posttests ε_j soient globalement très corrélées entre elles.
En particulier, la positivité de τ sera assurée quand $\beta_{YX} > 0$ et $\rho_{\delta\varepsilon} \leq 0$, ou bien quand $\beta_{YX} > 1$ et les posttests sont au moins aussi fidèles que les prétests ($\lambda_y \geq \lambda_x$).

A présent, revenons à notre contexte où le traitement est une variable observée qui s'adresse de manière privilégiée à des sujets de compétence extrême. Il s'agit d'un enrichissement qui s'adresse aux plus forts, ce qui implique certainement ρ_{XT} positif, ou au contraire, d'une remédiation qui s'adresse aux plus faibles, ce qui implique sûrement ρ_{XT} négatif.

A priori, β_{yT} devrait pouvoir de manière égale sous-estimer ou surestimer β_{YT}. Or la symétrie théorique des possibilités est maintenant rompue en acceptant l'hypothèse très plausible que $\tau > 0$: α a le signe de ρ_{XT} et on peut avancer raisonnablement que seront plus fréquentes dans la réalité les situations suivantes.

1) Quand T est un *enrichissement,* $\alpha > 0$.

Quand β_{YT} est positif, β_{yT} maintient cette positivité. Quand β_{YT} est négatif ou nul, β_{yT} le surévalue.

2) Quand T est une *remédiation*, $\alpha < 0$.

Quand β_{YT} est positif ou nul, β_{yT} le sous-évalue. Quand β_{YT} est négatif, β_{yT} maintient cette négativité.

Ceci *démontre* un phénomène pervers relevé empiriquement depuis quelque temps : l'interprétation de β_{yT} tend vers une conclusion partiale, optimiste pour un affermissement, pessimiste pour une remédiation.

"The point to be emphasized here is that the use of standard regression approaches often results in unwarranted conclusions of harmful treatment effects when "compensatory" treatments are evaluated where the persons receiving the treatment do less well than controls at the pretest. Similarly, they often lead to unwarranted conclusions of beneficial impact when programs are evaluated where the initially more fortunate group has greater exposure to the treatment..."
[Cook, Campbell, p. 301]

Quand le traitement est une répartition en filières fortes ou faibles, ce phénomène est également relevé dans [Slavin].

5. ESTIMATION DU COEFFICIENT NET

Ce qui précède montre assez qu'on ne peut se satisfaire du coefficient brut β_{yT} et qu'il faut pousser l'analyse jusqu'au coefficient net β_{YT}.

Le modèle présenté ci-avant est un modèle de structure de covariances [Bollen, chap. 8] car il combine

- des mesures des variables latentes exogènes X et T par les équations (2) et l'identité

$$T = \lambda_T T + \delta_T \qquad \lambda_T = 1 \qquad \mathrm{var}(\delta_T) = 0$$

qui exprime que T est à la fois variable observable et facteur;

- des mesures d'une variable latente endogène Y par les équations (3) ;

- une structure causale linéaire entre X, T, Y par l'équation (1).

Dans ce modèle les paramètres à estimer sont les saturations λ_i et λ_j , les covariances et variances d'erreurs

$$\theta_{ik} = \mathrm{cov}(\delta_i,\delta_k) \qquad \theta_{jk}=\mathrm{cov}(\varepsilon_j,\delta_k) \qquad \theta_{jl}=\mathrm{cov}(\varepsilon_j,\varepsilon_l)$$

la corrélation ρ_{XT}, les coefficients β_{YT} et β_{YX}, la variance résiduelle $\mathrm{var}(\zeta)$. Des logiciels tirent les estimations de ces paramètres de la matrice des covariances entre variables observées T, x_i, y_j. Différentes méthodes d'estimation (moindres carrés, moindres carrés pondérés, maximum de vraisemblance, etc) sont disponibles et les programmes permettent de confronter le modèle aux données par des épreuves d'ajustement (chi carré).

Voici un exemple d'application de cette procédure [Meuret].

L'échantillon est constitué de quelques milliers d'élèves suivis de 6e en 5e de collèges. Le traitement T est l'appartenance à une ZEP (zone d'éducation prioritaire); la variable exogène X est une compétence scolaire initiale PRE mesurée par 4 prétests en français et 4 prétests en mathématique ; la variable endogène Y est la compétence scolaire finale POST mesurée par 2 posttests en français et 2 posttests en mathématique. Les erreurs de mesure des tests de même nature pouvaient être corrélées, y compris entre prétests et posttests. En outre sont intervenues comme covariables explicatives de POST la catégorie socio-professionnelle des parents, la nationalité et l'âge, variables observées identifiées à des facteurs car supposées être mesurées sans erreur.

Le modèle a été traité par le logiciel LISREL [Jöreskog, Sörbom]. Les coefficients β_{YT} et β_{YX} ont été estimés à -0.042 et 0.940. Le coefficient brut β_{yT} estimé dans une régression multiple "naïve" était - 0.130.

6. IDENTIFICATION DU MODÈLE

Subsiste la question de la spécification du modèle. De ce point de vue (mais nullement de celui de l'estimation), l'identification des paramètres β_{YT}, β_{YX} est visiblement équivalente par les équations (11) à celle des paramètres ρ_{YT}, ρ_{XY}.

Les équations reliant les quantités observables aux paramètres sont alors les équations (7) et (9) $\qquad \rho_{iT} = \lambda_i \rho_{XT} \qquad \rho_{ik} = \lambda_i \lambda_k + \theta_{ik}$

et les suivantes déduites immédiatement des équations (2) et (3)

$$\mathrm{cov}(y_j,T) = \rho_{jT} = \lambda_j \rho_{YT} \qquad \mathrm{cov}(y_j,x_k) = \rho_{jk} = \lambda_j \lambda_k \rho_{XY} + \theta_{jk} \qquad \mathrm{cov}(y_j,y_l) = \rho_{jl} = \lambda_j \lambda_l + \theta_{jl}$$

On reconnaît dans ce système l'équation d'une analyse factorielle confirmatoire [Bollen, equ. (7.5)] $\qquad \Sigma=\Lambda\Phi\Lambda'+\Theta \qquad$ où

$$\Sigma = \begin{array}{cc} \left[\begin{array}{ccc} 1 & & \\ \rho_{iT} & \rho_{ik} & \\ \rho_{jT} & \rho_{jk} & \rho_{jl} \end{array}\right] & \begin{array}{c} T \\ x_i \\ y_j \end{array} \\ \begin{array}{ccc} T & x_k & y_l \end{array} & \end{array} \qquad \Lambda = \begin{array}{cc} \left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & \lambda_i & 0 \\ 0 & 0 & \lambda_j \end{array}\right] & \begin{array}{c} T \\ x_i \\ y_j \end{array} \\ \begin{array}{ccc} T & X & Y \end{array} & \end{array}$$

$$\Phi = \begin{array}{cc} \left[\begin{array}{ccc} 1 & & \\ \rho_{XT} & 1 & \\ \rho_{YT} & \rho_{YX} & 1 \end{array}\right] & \begin{array}{c} T \\ X \\ Y \end{array} \\ \begin{array}{ccc} T & X & Y \end{array} & \end{array} \qquad \Theta = \begin{array}{cc} \left[\begin{array}{ccc} 0 & & \\ 0 & \theta_{ik} & \\ 0 & \theta_{jk} & \theta_{jl} \end{array}\right] & \begin{array}{c} \\ \delta_i \\ \varepsilon_j \end{array} \\ \begin{array}{ccc} & \delta_k & \varepsilon_l \end{array} & \end{array}$$

Dans cette généralité, le modèle est sous-identifié. Il n'est donc estimable et falsifiable qu'en ajoutant des contraintes a priori sur ses paramètres.

ρ_{XT}, ρ_{YT} (ou β_{YT}) et ρ_{XY} (ou β_{YX}) doivent évidemment rester libres car ils marquent les relations interrogées entre T, X et Y. Puisque

$$1=\rho_{ii}=\lambda_i^2+\theta_{ii} \qquad\qquad 1=\rho_{jj}=\lambda_j^2+\theta_{jj}$$

contraindre les fidélités λ_i^2 ou λ_j^2 des mesures x_i ou y_j revient à contraindre les variances θ_{ii} ou θ_{jj}. Les covariances d'erreurs de mesure θ_{ik}, θ_{jk}, θ_{jl} sont donc les seuls paramètres auxquels des liaisons puissent être décemment imposées. Celles-ci seront communément des fixations à une valeur a priori, le plus souvent 0.

Dans ce cadre pratique (ne fixer que des covariances d'erreurs), montrons que pour que le modèle soit identifié il faut et il suffit qu'au moins un $\theta_{ik}=\text{cov}(\delta_i,\delta_k)$ avec $\rho_{iT}\rho_{kT}\neq 0$, au moins un $\theta_{jk}=\text{cov}(\varepsilon_j,\delta_k)$ avec $\rho_{jT}\ \rho_{kT}\neq 0$ et au moins un $\theta_{jl}=\text{cov}(\varepsilon_j,\varepsilon_l)$ avec $\rho_{jT}\ \rho_{lT}\neq 0$ soient fixés.

Si tout θ_{ik} est laissé libre, fixer certains des θ_{jk}, θ_{jl}, voire tous, n'empêche pas que les transformations

$$\lambda_i \to a\lambda_i \qquad \rho_{XT}\to\rho_{XT}/a \qquad \theta_{ik}\to\theta_{ik}+\lambda_i\,\lambda_k(1-a^2) \qquad \rho_{XY}\to\rho_{XY}/a$$

où $a\neq 0,\pm 1$ laissent invariantes les équations

$\rho_{iT}=\lambda_i\rho_{XT}=(a\lambda_i)(\rho_{XT}/a)$

$\rho_{ik}=\lambda_i\lambda_k+\theta_{ik}=(a\lambda_i)(a\lambda_k)+(\theta_{ik}+\lambda_i\lambda_k(1-a^2))$

$\rho_{jk}=\lambda_j\lambda_k\rho_{XY}+\theta_{jk}=\lambda_j(a\lambda_k)\ (\rho_{XY}/a)+\theta_{jk}$

Si tout θ_{jk} est laissé libre, fixer certains des θ_{ik},θ_{jl}, voire tous, n'empêche pas que les transformations

$$\rho_{XY}\to\rho_{XY}+a \qquad \theta_{jk}\to\theta_{jk}-a\lambda_j\,\lambda_k$$

où $a\neq 0$ laissent invariantes les équations

$$\rho_{jk}=\lambda_j\lambda_k\rho_{XY}+\theta_{jk}=\lambda_j\lambda_k(\rho_{XY}+a)+(\theta_{jk}-a\lambda_j\lambda_k)$$

Si tout θ_{jl} est laissé libre, fixer certains des θ_{ik}, θ_{jk}, voire tous, n'empêche pas que les transformations

$$\lambda_j\to a\lambda_j \qquad \rho_{YT}\to\rho_{YT}/a \qquad \theta_{jl}\to\theta_{jl}+\lambda_j\,\lambda_l(1-a^2) \qquad \rho_{XY}\to\rho_{XY}/a$$

où $a\neq 0,\pm 1$ laissent invariantes les équations.

Réciproquement la fixation d'un θ_{ik}, d'un θ_{jk} et d'un θ_{jl} permet de tirer une valeur définie pour les autres paramètres.

Pour $\theta_{12}=\text{cov}(\delta_1,\delta_2)$, supposons que $\rho_{1T}\rho_{2T}\neq 0$. Une valeur fixée de θ_{12} est choisie de telle sorte que

$$0<\frac{\rho_{1T}\rho_{2T}}{\rho_{12}-\theta_{12}}<1$$

$\rho_{1T}=\lambda_1\rho_{XT} \qquad \rho_{2T}=\lambda_2\rho_{XT} \qquad \rho_{12}=\lambda_1\lambda_2+\theta_{12}$ donnent

$$\rho_{XT}=\sqrt{\frac{\rho_{1T}\,\rho_{2T}}{\rho_{12}-\theta_{12}}}\neq 0$$

On obtient alors $\lambda_i=\rho_{iT}/\rho_{XT}$,ensuite les autres $\theta_{ik}=\rho_{ik}-\lambda_i\lambda_k$
Avec $\rho_{1T}\neq 0$, une autre option est de fixer $\theta_{11}=var(\delta_1)$ dans $]0,1[$ qui livrera $\lambda_1=\sqrt{1-\theta_{11}}$, $\rho_{XT}=\rho_{1T}/\lambda_1$, le reste étant identique.

Parallèlement la fixation d'un $\theta_{jl}=cov(\varepsilon_j,\varepsilon_l)$ avec $\rho_{jT}\rho_{lT}\neq 0$ suffit à fournir successivement $\rho_{YT}\neq 0$, les λ_j, les autres θ_{jl}.

Enfin on peut fixer $\theta_{32}=cov(\varepsilon_3,\delta_2)$ quand on a trouvé un y_3 et un x_2 tels que $\rho_{3T}\rho_{2T}\neq 0$. Les démarches précédentes ont déjà donné les λ_j dont $\lambda_3=\rho_{3T}/\rho_{YT}\neq 0$, et les λ_k dont $\lambda_2=\rho_{2T}/\rho_{XT}\neq 0$. Il vient finalement $\rho_{XY}=\dfrac{\rho_{32}-\theta_{32}}{\lambda_3\lambda_2}$ puis les autres $\theta_{jk}=\rho_{jk}-\lambda_j\lambda_k\rho_{XY}$.

7. IMPORTANCE DE LA SPÉCIFICATION DU MODÈLE

Ainsi tout ensemble de données observées laisse au moins trois degrés de liberté à la spécification pour engendrer divers modèles identifiés possibles qui auront chacun leur propre coefficient net β_{YT}.

Par (11)

$$\beta_{YT}=\frac{\rho_{YT}-\rho_{XY}\rho_{XT}}{1-\rho_{XT}^2} \qquad (14)$$

est fonction des corrélations entre variables fondamentales. Un point important de la procédure empirique sera le choix <u>préalable</u> de liaisons entre certains θ_{ik}, θ_{jk}, θ_{jl} puisque ce choix conditionne en bout de course la valeur estimée de β_{YT}.

L'exemple suivant est édifiant à cet égard. Supposons que les observations aient généré la matrice suivante des corrélations

$$\begin{array}{cc} \begin{bmatrix} 1 & & & \\ 0.792 & 1 & & \\ 0.792 & 0.64 & 1 & \\ 0 & 0 & 0 & 1 \end{bmatrix} & \begin{matrix} T \\ x_1 \\ x_2 \\ y_3 \end{matrix} \end{array}$$

Ici Y est mesuré par un seul indicateur y_3. La variable y construite par (4) est forcément y_3. Comme

$$\rho_{yT}=\rho_{3T}=0 \qquad \vec{\rho}_{xy}=(\rho_{13},\rho_{23})=(0,0)$$

(6) donne le coefficient brut $\beta_{yT}=0$.

Pour identifier le modèle, spécifions d'abord que θ_{21}, covariance des erreurs de mesure de x_2 et x_1, soit nulle. Les cinq équations

$$\rho_{1T}=0.792=\lambda_1\rho_{XT} \qquad \rho_{11}=1=\lambda_1^2+\theta_{11} \qquad \rho_{2T}=0.792=\lambda_2\rho_{XT}$$
$$\rho_{21}=0.64=\lambda_2\lambda_1+\theta_{21}=\lambda_2\lambda_1 \qquad \rho_{22}=1=\lambda_2^2+\theta_{22}$$

fournissent

$$\rho_{XT}=\sqrt{\rho_{1T}\rho_{2T}/\rho_{21}}=0.99 \quad \lambda_1=\lambda_2=0.8 \qquad \theta_{11}=\theta_{22}=0.36$$

En fixant ensuite θ_{33}

$$\rho_{33} = 1 = \lambda_3^2 + \theta_{33} \quad \text{donne} \quad \lambda_3 = \sqrt{1 - \theta_{33}}$$

$$\rho_{3T} = 0 = \lambda_3 \rho_{YT} \quad \text{donne} \quad \rho_{YT} = 0$$

Une fois que θ_{31} a été fixé, les deux dernières équations

$$\rho_{31} = 0 = \lambda_3 \lambda_1 \rho_{XY} + \theta_{31} \qquad \rho_{32} = 0 = \lambda_3 \lambda_2 \rho_{XY} + \theta_{32} \qquad \text{c'est-à-dire}$$

$$\theta_{31} = \theta_{32} \qquad \theta_{31}^2 < \inf\{0.64(1 - 0.99^2)\,(1 - \theta_{33}),\ 0.18\theta_{33}\}$$

livrent

$$\rho_{XY} = -\,\theta_{31} / (0.8\sqrt{1 - \theta_{33}})\,, \quad \text{d'où par (14)}$$

$$\beta_{YT} = \frac{0.99\theta_{31}}{0.8(1-0.99^2)\sqrt{1-\theta_{33}}}$$

Les paramètres θ_{33} et θ_{31} sont spécifiés arbitrairement à ceci près qu'ils doivent préserver la définie positivité des matrices

$$\begin{bmatrix} 1 & & \\ \rho_{XT} & 1 & \\ \rho_{YT} & \rho_{YX} & 1 \end{bmatrix} = \begin{bmatrix} 1 & & \\ 0.99 & 1 & \\ 0 & \dfrac{-\theta_{31}}{0.8\sqrt{1-\theta_{33}}} & 1 \end{bmatrix}$$

et

$$\begin{bmatrix} \theta_{11} & & \\ \theta_{21} & \theta_{22} & \\ \theta_{31} & \theta_{32} & \theta_{33} \end{bmatrix} = \begin{bmatrix} 0.36 & & \\ 0 & 0.36 & \\ \theta_{31} & \theta_{31} & \theta_{33} \end{bmatrix}$$

On vérifie aisément qu'il faut et il suffit pour cela que

$$\theta_{31}^2 < \inf\{0.64(1 - 0.99^2)\,(1 - \theta_{33}),\ 0.18\theta_{33}\}$$

condition qui, pour $\theta_{33} > 0.066$, autorise $|\theta_{31}| / \sqrt{1-\theta_{33}} < 0.8\sqrt{1-0.99^2}$

et donc des $|\beta_{YT}|$ aussi grands que $0.99 / \sqrt{1-0.99^2} = 7.01$

Par exemple, avec $\theta_{33} = 0.1$, on peut prendre $\theta_{31} = \pm\, 0.1$ qui donne $\beta_{YT} = \pm\, 6.55$

Le contraste possible entre $\beta_{YT} = 0$ et β_{YT} vient de la forte corrélation entre compétence initiale et traitement ($\rho_{XT} = 0.99$) conjuguée à une corrélation entre compétences initiale et finale masquée dans les observations ($\rho_{31} = 0 = 0.8\lambda_3 \rho_{XY} + \theta_{31}$) par l'association des erreurs de leurs mesures (θ_{31}).

La leçon de ce petit exemple est dure : avec des données plausibles qui engendrent un coefficient brut nul, des modèles compatibles, exactement ajustés, peuvent fournir des coefficients nets non nuls extrêmement différents <u>suivant</u> la spécification imposée aux erreurs de mesure (ci-dessus par le canal de $\theta_{21},\theta_{33},\theta_{31}$).

8. EXTENSION À D'AUTRES COVARIABLES

Il peut y avoir des variables autres que T et X qui influencent Y. Des exemples ont été cités dans le paragraphe 5. Tenir compte du vecteur $\vec{R}$ de ces covariables supplémentaires complique diablement les choses quand $\vec{R}$ est imprécisément mesuré.

On peut montrer que β_{yT} se présente encore comme somme de $\lambda_y\beta_{YT}$ et d'un biais additif qui est le produit scalaire de $(\rho_{XT},\vec{\rho}_{RT})$ par un vecteur contenant tous les paramètres de mesure et les coefficients de Y sur X et $\vec{R}$. Le signe de ce biais devient totalement imprévisible.

Mais supposons seulement que les covariables supplémentaires de $\vec{R}$ soient <u>mesurées sans erreur</u>. Regroupant dans le vecteur $\vec{C}=(T,\vec{R})$ le traitement et ces variables nouvelles, simultanément latentes et observées, les deux équations de régression à comparer, correspondant à (1) et (5), sont

$$Y = \vec{\beta}_{YC}'\,\vec{C} + \beta_{YX}\,X + \zeta \qquad\qquad y = \vec{\beta}_{yC}'\,\vec{C} + \sum_i \beta_{yi}\,x_i + \zeta'$$

Conduisons alors les calculs comme dans le paragraphe 3 en lisant systématiquement T comme le vecteur $\vec{C}$. A (6) correspond

$$\vec{\beta}_{yC} = (P_{CC} - P_{xC}'\,\Sigma_{xx}^{-1}\,P_{xC})^{-1}\,(\vec{\rho}_{yC} - P_{xC}'\,\Sigma_{xx}^{-1}\,\vec{\rho}_{xy}) \qquad\qquad \text{où}$$

P_{CC} est la matrice des corrélations à l'intérieur de $\vec{C}$,
P_{xC} est la matrice des corrélations entre $\vec{x}$ et $\vec{C}$,
$\vec{\rho}_{yc}$ est le vecteur des corrélations entre y et $\vec{C}$.

Tenant compte de

$P_{xC} = \vec{\lambda}\,\vec{\rho}_{XC}'$ où $\vec{\rho}_{XC}$ est le vecteur des corrélations entre X et $\vec{C}$,
$\vec{\rho}_{yC} = \lambda_y\,\vec{\rho}_{YC}$ où $\vec{\rho}_{YC}$ est le vecteur des corrélations entre Y et $\vec{C}$

et des équations normales correspondant à (11)

$$\vec{\rho}_{YC} = P_{CC}\,\vec{\beta}_{YC} + \beta_{YX}\,\vec{\rho}_{XC} \qquad\qquad \rho_{XY} = \vec{\beta}_{YC}'\,\vec{\rho}_{XC} + \beta_{YX}$$

on obtient la version plus générale de (12)

$$\vec{\beta}_{yC} = \lambda_y\,\vec{\beta}_{YC} + \frac{\tau P_{CC}^{-1}\,\vec{\rho}_{XC}}{1 - \lambda_x^2\,\vec{\rho}_{XC}'\,P_{CC}^{-1}\,\vec{\rho}_{XC}}$$

où le dénominateur est toujours positif et où τ est la quantité critique discutée dans le paragraphe 4.

$$P_{CC} = \begin{bmatrix} 1 & \vec{\rho}_{RT}' \\ \vec{\rho}_{RT} & P_{RR} \end{bmatrix} \qquad\qquad \vec{\rho}_{XC} = \begin{bmatrix} \rho_{XT} \\ \vec{\rho}_{XR} \end{bmatrix}$$

entraînent pour la première composante de l'égalité ci-avant

$$\beta_{yT} = \lambda_y \, \beta_{YT} + \frac{\tau(\rho_{XT} - \vec{\rho}\,'_{XR} \, P_{RR}^{-1} \, \vec{\rho}_{RT})}{(1 - \lambda_x^2 \, \vec{\rho}\,'_{XC} \, P_{CC}^{-1} \, \vec{\rho}_{XC}) \, (1 - \vec{\rho}\,'_{RT} \, P_{RR}^{-1} \, \vec{\rho}_{RT})}$$

Admettant encore que τ est positif, le signe du biais additif est celui de la corrélation *partielle* de X et T avec neutralisation des variables de $\vec{R}$:

$$\rho_{XT.R} = \frac{\rho_{XT} - \vec{\rho}\,'_{XR} \, P_{RR}^{-1} \, \vec{\rho}_{RT}}{\sqrt{(1 - \vec{\rho}\,'_{XR} \, P_{RR}^{-1} \, \vec{\rho}_{XR}) \, (1 - \vec{\rho}\,'_{RT} \, P_{RR}^{-1} \, \vec{\rho}_{RT})}}$$

Ainsi, si toutes les covariables autres que la compétence initiale sont exactement mesurées, la thèse précédente peut être encore soutenue.

9. CONCLUSION

Un phénomène pervers, observé depuis longtemps, est ici mathématiquement mis en évidence. Si l'effet réel d'une remédiation (resp. enrichissement) est nocif (resp. bénéfique), le sens de l'effet apparent est conservé tel quel; par contre, si l'effet réel d'une remédiation (resp. enrichissement) est bénéfique (resp. nocif), l'effet apparent est atténué, et son sens peut être renversé.

Cette perversité due aux erreurs de mesure ne peut être corrigée que par une spécification correcte d'un modèle fondamental où jouent deux variables latentes au moins, les compétences initiale et finale "vraies". Comme celui-ci est chroniquement sous-identifié, cela exige de lui imposer des contraintes a priori. La difficulté est déplacée au plan théorique.

Le chercheur devra, préalablement aux observations et aux calculs, décider sur des bases uniquement substantielles que telles erreurs des mesures ont telles covariances. Entreprise responsable et délicate car elle est lourde de conséquences sur l'évaluation de l'efficacité "vraie".

Il reste à escompter que les indicateurs x_i, y_j des compétences initiale et finale soient assez familiers pour pouvoir y détecter incontestablement des erreurs δ_i, ε_j non corrélées. A défaut d'un ensemble suffisant de telles informations, et seulement à titre exploratoire, la moins mauvaise stratégie semble de partir du modèle le plus serré où toutes les corrélations d'erreurs soient d'office annulées. Les indices de modification [Bollen, p. 299] pourront alors guider un relâchement de paramètres injustement fixés, en vue d'améliorer l'adéquation du modèle.

BIBLIOGRAPHIE

BOLLEN K. A., *Structural equations with latent variables*, New York, Wiley, 1989.

COOK T. D., CAMPBELL D. T., *Quasi-experimentation, design and analysis issues for field settings*, Chicago, Rand Mc Nally, 1979.

JÖRESKOG K. G., SÖRBOM D., LISREL 7 : *A guide to the program and applications*, Chicago, SPSS Inc., 1989.

MEURET D., L'efficacité de la politique des zones d'éducation prioritaires dans les collèges, *Revue française de pédagogie*, à paraître.

MINGAT A., "Les activités de rééducation GAPP à l'école primaire, analyse du fonctionnement et évaluation des effets", *Revue française de sociologie*, vol. 32 (1991), n° 4.

Le Monde de l'Education, "Courrier des lecteurs", nos 180 et 181, mars et avril 1991.

SLAVIN R. E., "Achievement effects of ability grouping in secondary schools : a best-evidence synthesis", *Review of educational research*, vol. 60 (1990), n°5, pp. 471-499.