

# DISEÑOS LONGITUDINALES EN PANEL: ALTERNATIVA DE ANALISIS DE DATOS MEDIANTE LOS SISTEMAS DE ECUACIONES ESTRUCTURALES

Jaune ARNAU GRAS y Joan GUARDIA OLMOS  
Facultat de Psicologia. Universitat de Barcelona.

## RESUMEN

El uso de los diseños longitudinales en panel para el estudio del efecto temporal en las relaciones entre las variables parece ser una alternativa válida para la obtención de evidencia empírica a propósito de la posible causalidad entre las variables definidas. Sin embargo, el análisis tradicional de las distintas correlaciones adolece de algunas deficiencias centradas fundamentalmente en la dificultad de recoger los auténticos efectos entre variables. En este trabajo se pretende presentar la traducción de los diferentes modelos estadísticos planteables mediante diseños de panel de dos tandas (2w2v) a sistemas de ecuaciones estructurales, así como las mejoras que este tipo de análisis supone con respecto al estudio de los distintos efectos propios de esta estrategia de investigación.

**Palabras clave:** Diseños longitudinales. Ecuaciones estructurales. Diseños en panel.

## ABSTRACT

*Longitudinal panel designs: alternative data analysis with structural equations.* - The use of the longitudinal panel design for the study of the temporal effect on the variables relationships, seems a valid alternative for to obtain empirical evidence about the hypothetical causal relation between variables. However, the traditional analysis of the different correlations present some deficiencies, essentially centered in the difficulty for to show the true effects between variables. The aim of this paper is to present the translation of the different statistical models referred with the panel design with two waves (2w2v); to the structural equations system and, also, the improvements of this kind of analysis, with respect of the study of the different effects of this procedure research.

**Key words:** Longitudinal designs. Structural equations. Panel designs.

## INTRODUCCION

El estudio de las distintas estrategias estadísticas para el análisis de las relaciones entre variables, contemplan con sumo cuidado la cuestión planteada cuando el

marco de recogida de datos es longitudinal. De hecho, los diseños longitudinales se presentan como uno de los procedimientos más ampliamente aplicados. En consecuencia, todo aquello que se conecte con las posibilidades de análisis de datos a partir de estos diseños es, efectivamente, relevante. La complejidad inhe-

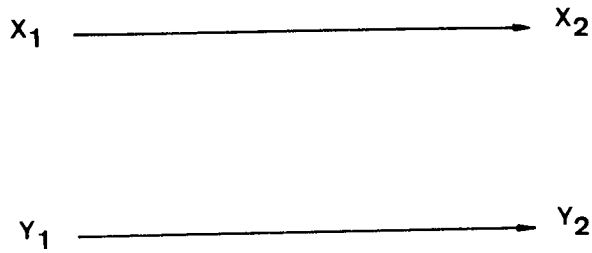


Figura nº 1.- Diseño en panel de dos medidas en dos tandas.

rente a estos diseños viene dada por la presencia de la temporalidad, lo que permite, a pesar de incrementar su dificultad, estudiar dos efectos globales entre variables: La relación transversal entre variables y la longitudinal (Rogosa, 1979).

En el presente trabajo se persigue plantear de forma breve la posibilidad de aplicación de los sistemas de ecuaciones estructurales en el ámbito de los diseños longitudinales, empleando el diseño básico dentro de los longitudinales: el diseño en panel con dos variables y dos medidas a través del tiempo (2W2V) (Figura 1). Se trata del diseño más básico de los de su género (Kenny, 1975). Fundamentalmente, permite efectuar diversos tipos de análisis. En primer lugar, se obtiene evidencia de los cambios a lo largo del tiempo, tomando cada atributo por separado. Como segundo nivel, se perfila la posibilidad de relacionar todas las observaciones por pares, y en último término, facilita el estudio de los denominados «efectos condicionales», caracterizados a partir de las correlaciones trans-retardadas. No se trata aquí de debatir las características del procedimiento de la correlación trans-retardada (Campbell y Stanley, 1966; Cook y Campbell, 1976); sino la de presentar una alternativa que consideramos más global y que permite un análisis exhaustivo de los distintos efectos que, la ya mencionada correlación no contempla (Rogosa, 1980). Por citar los dos elementos definitorios de

la aportación de los modelos estructurales es preciso pensar que esta técnica supone, por una parte, la inclusión de efectos espurios en el diseño; y por otra, el análisis de los errores de medida. A este respecto, consideramos que la evaluación empírica del error de medida es más exhaustiva que lo que se traduce en su repercusión mediante la estrategia de la «desatenuación» en la estimación de parámetros.

Así, el empleo de la regresión o del «Path Analysis» permite superar algunas de las concepciones limitadas de las correlaciones trans-retardadas, pero ambas quedan englobadas en los sistemas de ecuaciones estructurales, a la vez que éstos incorporan los modelos de medida, con un estudio más exhaustivo del cambio entre momentos temporales (Jöreskog y Sörbom, 1977).

#### MODELO LINEAL DE LA REGRESION

El empleo del modelo lineal de la regresión en este tipo de diseño se establece a partir de la consideración de dos modelos para la evaluación del cambio: modelos condicionales e incondicionales.

El primero de ellos se caracteriza por tomar como fijos los valores de la primera medida (X) y opera con la distribución de la segunda medida (Y) en base a los valo-

res de la primera. La ecuación general que reproduce esta situación es:

$$Y = a + \beta X + \varepsilon \quad (1)$$

La operativa, en este caso, es harto conocida y no es necesario efectuar un análisis exhaustivo de la misma (Gujarati, 1988). En el caso de los modelos incondicionales, el cambio se evalúa en base a diferencias netas entre observaciones ( $d = x_2 - x_1$ ). Existen algunas dudas (Plewis, 1985) acerca de la adecuación de este tipo de modelos, puesto que ignora la dirección del cambio. A pesar de ello, la estrategia combinada de los modelos condicionales de reversión temporal, parece una solución de compromiso entre ambas opciones. De este modo, se puede plantear este último modelo de la siguiente forma:

$$X_t = \tau + \delta Y + \varepsilon \quad (2)$$

De esta forma, los cambios individuales evaluados en el modelo incondicional, se tratan en base a la dirección ( $x_1 \rightarrow y$ ) más que en sentido contrario, lo cual se adecúa perfectamente a los requerimientos del modelo condicional.

Esta breve referencia al empleo del modelo de la regresión debe verse amplia-

da (Plewis, 1985) con el análisis de las situaciones en las que el cambio se evalúa a partir de modelos individuales para cada par longitudinal de observaciones. En efecto, parecen claramente diferenciables las situaciones que se caracterizan por obtener combinaciones lineales, para cada par, que sean paralelas, divergentes o convergentes. Cada uno de esos resultados merecen un estudio particular. Sin embargo, debemos recordar el ya conocido fenómeno de la escasa diferenciación entre regresiones, a pesar de que los centroides de cada uno de los grupos de observación sean netamente diferentes. El siguiente gráfico muestra esta situación conocida como la paradoja de Lord.

El cambio observado entre ambas expresiones viene determinado por el valor de  $\beta_0$ . Su análisis indicaría un cambio de nivel escaso, en tanto en cuanto los centroides muestran una distancia entre sí mucho más considerable.

Por último, debe tenerse en cuenta que, el modelo lineal de la regresión se constituye en el común denominador de las distintas técnicas que a continuación se citarán, y en consecuencia deberán tratarse como extensiones del citado modelo.

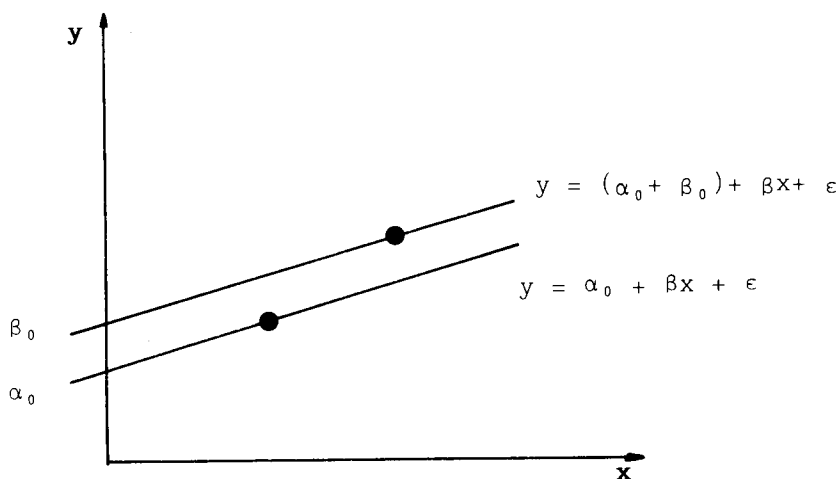


Figura nº 2.- Ejemplo gráfico de la paradoja de Lord.

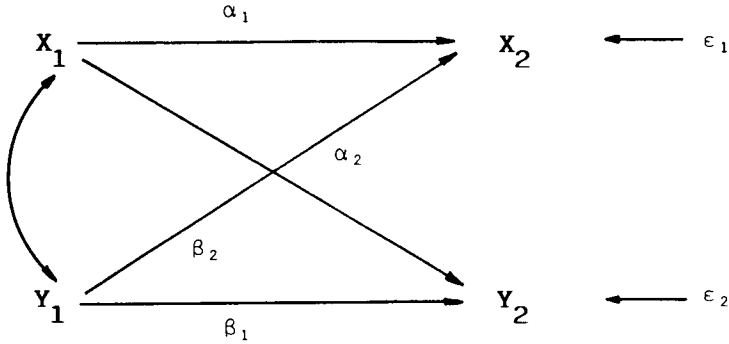


Figura nº 3.- Path Diagrama Diseño en Panel.

TECNICA DEL «PATH ANALYSIS»

1.- Diseño en panel simple (2W2V) para el estudio de la dirección de la relación entre las variables

Es este el caso que supone la aportación, por lo que se refiere al análisis de datos, más clara con respecto al método de Campbell de la Correlación trans-retardada.

En concreto, en la figura siguiente se plantea el «path diagrama» correspondiente a un diseño en panel con dos variables evaluadas en dos momentos:

Por lo que se refiere a los efectos cruzados, constituyen el punto de interés a tener presente, puesto que nos servirán para obtener evidencia empírica que permita establecer alguna hipótesis en torno a la dirección de la relación entre las variables implicadas. La estrategia de poner en práctica la técnica del Path Analysis no supone, en modo alguno, un aspecto diferenciador de lo que puede obtenerse del empleo del modelo de la regresión. El elemento importante que cabe destacar es el hecho de que, en lugar de contrastar directamente los coeficientes de correlación trans-retardados, evaluaremos estadísticamente las distintas cargas entre variables, para así determinar la posible direccionalidad. El anterior Path puede ex-

presarse mediante las siguientes expresiones:

$$X_2 = a_0 + a_1 x_1 + a_2 Y_1 + \epsilon_1 \quad (3)$$

$$Y_2 = \beta_0 + \beta_1 Y_1 + \beta_2 X_1 + \epsilon_2 \quad (4)$$

En consecuencia, será necesario obtener la estimación de los coeficientes  $a_2$  y  $\beta_2$  para la evaluación de cuál es el efecto más relevante. A la vista del diagrama, los valores estimados de  $a_1$  y  $\beta_1$  corresponden a la evaluación de la autocorrelación, que no son relevantes a la cuestión de la direccionalidad.

Parece complementario establecer la estimación de los parámetros de las dos ecuaciones, mediante el procedimiento de la regresión a través de OLS o mediante el método de los momentos propio del Path Analysis. Esta última no es más que una extensión de la regresión. Sin embargo cabe señalar dos aspectos importantes a este respecto (Plewis, 1985):

1.- La estimación mediante el método de los momentos supone una estimación global, en la que los parámetros se pueden expresar como una combinación lineal del resto. La modificación de un parámetro afecta al valor del resto.

2.- El método de los momentos permite descomponer los coeficientes de covarianza/correlación iniciales de forma exhaustiva, en base a los «recorridos» es-

tablecidos, en tanto que la regresión estima parámetros para cada ecuación separadamente.

De todas formas, es bien sabido que en determinadas circunstancias (tamaño de muestra, distribución de las variables, robustez...) ambos sistemas de estimación coinciden en las soluciones que ofrecen (Guàrdia, 1986).

Las expresiones anteriores, son susceptibles, obviamente, de ser sometidas al proceso típico de la estandarización para solventar problemas de escalas de medida. Con ello simplemente, obtendremos expresiones lineales a través del origen de coordenadas, al no existir punto de corte o constante. El modelo presentado plantea un inconveniente, propio de este tipo de tratamiento. En concreto, no permite el establecimiento de evidencia empírica del cambio operado entre ambas medidas. Caso de que decidiéramos que la segunda medida ha supuesto un cambio significativo en las medidas de las variables, éste podría expresarse del siguiente modo:

$$\begin{aligned} X_2 - X_1 &= a_0 + a_1 X_1 + a_2 Y_1 + \varepsilon_1 - X_1 = \\ &= X_1 (a_1 - 1) + (a_0 + a_2 Y_1 + \varepsilon_1) = \\ &= X_1 (a_1 - 1) + \Phi_1 \end{aligned} \quad (5)$$

donde  $\Phi_1$  depende del valor de  $Y_1$  y  $\varepsilon_1$ . Por su parte y de forma análoga, la expresión  $(Y_2 - Y_1)$  podría establecerse del siguiente modo:

$$Y_2 - Y_1 = Y_1 (\beta_1 - 1) + \Phi_2 \quad (6)$$

donde  $\Phi_2$  depende del valor de  $X_1$  y  $\varepsilon_2$ . A la vista de las anteriores expresiones, debe destacarse que el cambio entre medidas dependen de variables ajenas a la medida del cambio en sí misma. Es decir, la obtención de parámetros estructurales del cambio depende de las variancias de  $\Phi_1$ , lo cual es lo mismo que decir que la evidencia del cambio está conectada con variables ajenas al propio cambio y con

ello no estamos aportando información relacionada con la posible direccionabilidad de la relación (Plewis, 1985).

Finalmente, por lo que se refiere a este caso concreto, señalar que la obtención de las estimaciones de  $a_2$  y  $\beta_2$  deberá ir acompañada por la consecución de su error estándar y proceder acto seguido a la significación de ambos. A la vista de los resultados que se obtengan, podrá dilucidarse la posible dirección de la relación entre las variables:

$$t\text{-value}(a_2) = a_2 / SE(a_2) \quad (7)$$

$$t\text{-value}(\beta_2) = \beta_2 / SE(\beta_2) \quad (8)$$

Caso de que  $t(a_2)$  sea significativamente distinto de 0 y  $t(\beta_2)=0$  podrá pensarse en la dirección  $Y \rightarrow X$ , mientras que en caso contrario, evidentemente, la dirección planteada sería la contraria.

Algunas de las consideraciones efectuadas en este caso son válidas para el resto de las situaciones, por lo que no se plantearán de nuevo. Sin embargo, como servirán de argumento de algunas de las matizaciones, será necesario, como mínimo, mencionarlas, para no establecer excesivos saltos en el desarrollo de los distintos modelos.

## 2.- Diseño en panel (2W2V) con efectos simultáneos

Al margen de algunos comentarios efectuados acerca del modelo anterior, cabe señalar que ese modelo no puede ser considerado como un modelo completo, entendiéndose como tal aquel modelo que permite analizar todos los diferentes efectos que pueden establecerse entre las variables implicadas. De ahí que el perfeccionamiento del modelo anterior pase por la especificación de los efectos que en la segunda medida puede establecerse entre las variables. Nos encontramos, pues, con el modelo de efectos recíprocos o simultáneos (Modelos no recursivos) que se

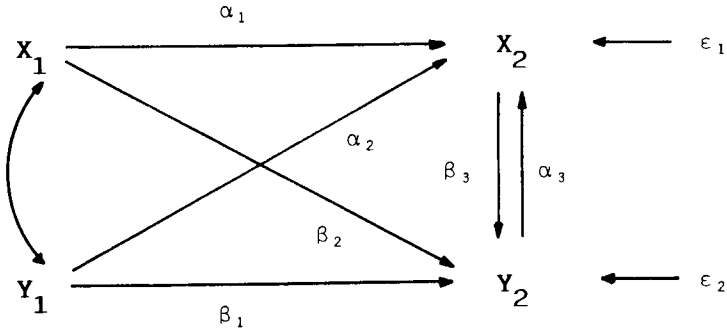


Figura nº 4.- Diseño en panel con efectos simultáneos.

representa mediante el siguiente Path Diagrama:

Este modelo es característico de los no recursivos y, en consecuencia, presenta todas las dificultades propias de estos modelos. Desde el punto de vista teórico, permite analizar dos efectos de carácter más exhaustivo que el modelo simple presentado anteriormente:

1.- Permite establecer efectos de carácter indirecto (modelo de feedback). Los efectos del tipo  $Y_1-X_2-Y_2$ , por un lado, y  $X_1-Y_2-X_2$ , por otro, no eran evaluables en el modelo simple, mientras que en el modelo ahora propuesto son perfectamente analizables. Los siguientes diagramas responden a este tipo de efectos (Rogosa, 1979).

2.- El establecimiento de efectos recíprocos supone una característica particular de los modelos de ecuaciones estructurales, puesto que la solución a la estimación

de parámetros de este tipo de modelos pasa necesariamente por el estudio exhaustivo de las condiciones de aplicación de los modelos estructurales. A pesar de que en sí mismos, los efectos recíprocos suscitan una gran controversia, en base a la verosimilitud de su especificación, no es éste el marco para tal discusión, a pesar de que no es gratuito plantearse su realidad teórica. Siguiendo el esquema anterior, las ecuaciones en las que se puede traducir el diagrama de efectos simultáneos son las siguientes:

$$X_2 = a_0 + a_1 X_1 + a_2 Y_1 + a_3 Y_2 + \epsilon_1 \quad (9)$$

$$Y_2 = \beta_0 + \beta_1 Y_1 + \beta_2 X_1 + \beta_3 X_2 + \epsilon_2 \quad (10)$$

En el modelo del apartado anterior el establecimiento de la direccionalidad de la relación pasaba por la comparación de las estimaciones de  $a_2$  y  $\beta_2$ . En este caso, el proceso es similar pero algo más amplio en lo que se refiere a los valores que de-

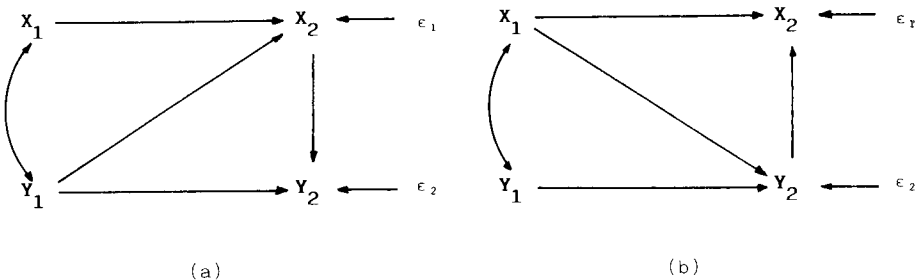


Figura nº 5.- Diagrama de los efectos analizables mediante el efecto simultáneo (a y b).

ben tenerse en cuenta. Así, puede resumirse que:

$$\text{Si } a_3 = 0 \text{ Y } \beta_3 / \neq 0 \implies X \text{ — } Y$$

$$\text{Si } \beta_3 = 0 \text{ Y } a_3 / \neq 0 \implies Y \text{ — } X$$

$$\text{Si } a_3 = \beta_3 = 0 \implies \text{la secuencia } Y_1 - X_2 - Y_2 \text{ puede ser plausible}$$

$$\text{Si } a_2 = \beta_3 = 0 \implies \text{la secuencia } X_1 - Y_2 - X_2 \text{ puede ser plausible}$$

Como decíamos anteriormente, este modelo es característico de los no recursivos, y presenta una serie de dificultades que pueden resumirse del siguiente modo (Berry, 1984):

1.- El modelo así planteado es un modelo no identificado. Es decir, es dudosa la obtención de soluciones únicas para cada una de las estimaciones de los parámetros. Para mostrar este aspecto, sólo cabe establecer los distintos elementos de las matrices fundamentales del modelo propuesto:

$$\text{Parámetros estructurales} = a_1 \ a_2 \ a_3 \ \beta_1 \ \beta_2 \ \beta_3$$

$$\text{Variancias perturbaciones} = E(\epsilon_1)^2 \ E(\epsilon_2)^2$$

$$\text{Covariancias exógenas} = E(X_1)^2 \ E(Y_1)^2 \ E(X_1, Y_1)$$

asumiendo los supuestos  $E(X_i) = 0$  y  $E(\epsilon_i) = 0$  propios del Path Analysis. Como se comprueba, se requieren once estimaciones, para lo cual se dispone de una matriz inicial de covarianzas/correlaciones de orden  $4 \times 4$ , lo que supone un total de 10 coeficientes iniciales para descomponer según el método de los momentos. En consecuencia, no es factible establecer ese sistema de ecuaciones puesto que se trata de un sistema sin solución única (hay más incógnitas que ecuaciones).

2.- Esta situación es la misma que se daba en el estudio del cambio ( $Y_2 - Y_1$ ) en el modelo anterior al ahora analizado. Como se decía, en su momento, este as-

pecto impedía el estudio del cambio en el diseño en panel simple. Ahora bien, en este caso la problemática es mayor, puesto que este hecho impediría el uso del modelo en su forma global.

Para abordar esta problemática, será necesario recurrir a las técnicas de reparametrización, usando la estrategia de la combinación lineal, propia de la condición de orden de los modelos estructurales. Así, será necesario eliminar el efecto simultáneo que se produce por el hecho de que  $X_2$  e  $Y_2$  aparezcan como «exógenas» en la ecuación complementaria a las que las definen como «endógenas». La combinación lineal de ambas ecuaciones se consigue substituyendo en cada ecuación el valor de  $X_2$  o  $Y_2$  por su combinación lineal. Así, pueden establecerse las siguientes expresiones:

Substituyendo la expresión para  $Y_2$  en la ecuación definitoria de  $X_2$  (fórmula 10 en 9) se obtiene:

$$X_2 = a_0 + a_1 X_1 + a_2 Y_1 + a_3 (\beta_0 + \beta_1 Y_1 + \beta_2 X_1 + \beta_3 X_2 + \epsilon_2) + \epsilon_1 \tag{12}$$

$$X_2 = a_0 + a_1 X_1 + a_2 Y_1 + a_3 \beta_0 + a_3 \beta_1 Y_1 + a_3 \beta_2 X_1 + a_3 \beta_3 X_2 + a_3 \epsilon_2 + \epsilon_1 \tag{13}$$

$$X_2 - (a_3 \beta_3 X_2) = (a_1 + a_3 \beta_2) X_1 + (a_2 + a_3 \beta_1) Y_1 + (a_0 + a_3 \beta_0) + (a_3 \epsilon_2 + \epsilon_1) \tag{14}$$

$$X_2 = (1 / (1 - a_3 \beta_3)) * [(a_0 + a_3 \beta_0) + a_1 + a_3 \beta_2) X_1 + (a_2 + a_3 \beta_1) Y_1 + (a_3 \epsilon_2 + \epsilon_1)] \tag{15}$$

Como se observa en esta última expresión, la variable  $X_2$  así definida no contempla a la variable  $Y_2$  en la derecha de la expresión. De hecho esta combinación lineal puede expresarse como:

$$X_2 = \delta * (\delta_0 + \delta_1 X_1 + \delta_2 Y_1 + \delta_3) \quad (16)$$

$$\text{siendo } \delta = [1 / (1 - a_3 \beta_3)] \quad (17)$$

$$\delta_0 = (a_0 + a_3 \beta_0)$$

$$\delta_1 = (a_1 + a_3 \beta_2)$$

$$\delta_2 = (a_2 + a_3 \beta_1)$$

$$\delta_3 = (a_3 \varepsilon_2 + \varepsilon_1)$$

ello supone reducir (estrategia de las formas reducidas) el número de variables endógenas que determinan a  $X_2$ , puesto que  $Y_2$  no aparece en el lado derecho de la última expresión. Análogamente, el mismo proceso aplicado a la ecuación definitoria de  $Y_2$  nos llevaría a la siguiente expresión final:

$$Y_2 = (1 / 1 - a_3 \beta_3) * [(\beta_0 + a_0 \beta_3) + (\beta_1 + a_2 \beta_3) Y_1 + (\beta_2 + a_1 \beta_3) X_1 + (\beta_3 \varepsilon_1 + \varepsilon_2)] \quad (18)$$

o la expresión equivalente:

$$Y_2 = \tau * (\tau_0 + \tau_1 Y_1 + \tau_2 X_1 + \tau_3) \quad (19)$$

que presenta sustituciones parecidas a las anteriormente especificadas para  $X_2$ . Igual que antes  $Y_2$  queda definida en base a variables diferentes a la otra endógena. Esta manipulación permite estudiar los efectos planteados inicialmente, dejando

de lado la problemática que supone la no identificación del modelo.

### 3.- Diseño en panel (2W2V) para la evaluación de efectos espurios

La segunda de las aportaciones de los modelos de ecuaciones estructurales (concretamente del Path Analysis y de su estimación de parámetros mediante el método de los momentos) estriba en la posibilidad de análisis de efectos espurios como procedimiento evaluativo de posibles efectos directos. De hecho, incluir efectos espurios en un diseño en panel permite concretar no sólo la dirección de la relación entre variables, sino si ésta está o no afectada por una tercera variable que sustenta la relación principal por analizar. Supone, en consecuencia, unir la posibilidad del diseño longitudinal con la introducción de la evaluación de los efectos espurios. A este respecto, al igual que se ha señalado en el diseño simple, el efecto de la espuricidad se puede plantear de dos formas distintas: La variable exógena que provoca el efecto espurio puede o no ser constante a través de las medidas. Estas dos posibilidades se reflejan en la figura número 6.

No se ha especificado en los dos diagramas anteriores ningún efecto simultáneo para no incorporar la problemática de la no identificación del modelo. Las ecuaciones que se desprenden de los dos mo-

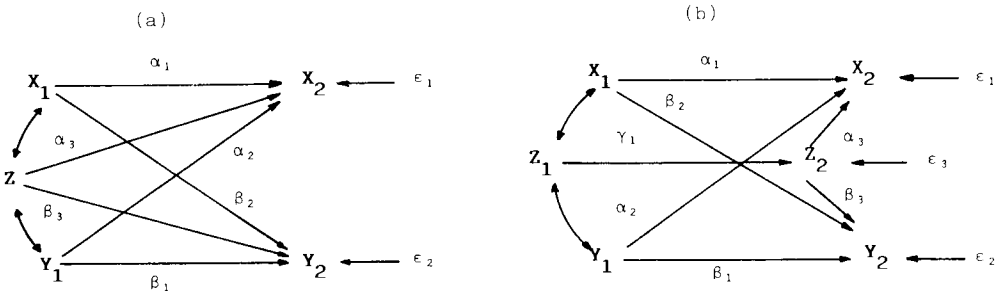


Figura nº 6.- Paths diagrams de los diseños en panel con una exógena constante (a) o con una exógena no constante (b).



delos se pueden expresar del modo planteado con posterioridad a la figura número 6.

Para el primer modelo (exógena constante)

$$X_2 = \alpha_0 + \alpha_1 X_1 + \alpha_2 Y_1 + \alpha_3 Z + \varepsilon_1 \quad (20)$$

$$Y_2 = \beta_0 + \beta_1 Y_1 + \beta_2 X_1 + \beta_3 Z + \varepsilon_2 \quad (21)$$

Las estimaciones de los parámetros  $\alpha_2$  y  $\beta_2$  se utilizarán para la evaluación de la direccionabilidad de la relación entre X e Y. Sin embargo, la presencia de Z provoca que los valores de  $\alpha_3$  y  $\beta_3$  sean imprescindibles para una correcta evaluación de la relación entre las variables. De hecho la similitud entre  $\alpha_3$  y  $\beta_3$  nos puede llevar a pensar en la posibilidad de que la relación representada por  $\alpha_2$  y  $\beta_2$  sea mantenida espúriamente por la presencia de Z. De la misma forma que se han comentado en el diseño simple, la estimación de esos parámetros puede efectuarse mediante los procedimientos propios del Path Analysis o de la regresión múltiple, considerando ambas ecuaciones por separado.

Para el segundo modelo (exógena no constante)

$$X_2 = \alpha_0 + \alpha_1 X_1 + \alpha_2 Y_1 + \alpha_3 Z_2 + \varepsilon_1 \quad (22)$$

$$Y_2 = \beta_0 + \beta_1 Y_1 + \beta_2 X_1 + \beta_3 Z_2 + \varepsilon_2 \quad (23)$$

$$Z_2 = \tau_0 + \tau_1 Z_1 + \varepsilon_3 \quad (24)$$

La incorporación de la variable  $Z_1$  provoca la configuración de una tercera ecuación estructural. En este caso la evaluación de la direccionabilidad de la espuricidad vendrá determinada por el estudio, al uso de lo visto anteriormente, a través de la obtención de las estimaciones de  $\alpha_2$ ,  $\beta_2$ ,  $\alpha_3$ ,  $\beta_3$  y  $\tau_2$ . A modo de ejemplo, la obtención de un valor de  $\alpha_3 > \alpha_2$  nos

llevaría al planteamiento de una relación entre Y y X mantenida por la presencia de  $Z_2$ . Análogamente, se pueden desarrollar todas las posibilidades correspondientes a los valores de los parámetros no relativos a la autocorrelación.

Piénsese, además, la complicación de análisis, pero también la exhaustividad del análisis, de cualquiera de los modelos anteriores a los que se les añada el efecto recíproco entre  $X_2$  e  $Y_2$ , a pesar de la dificultad de la especificación.

Este aspecto puede comprobarse con relativa facilidad si se plantean las ecuaciones que se originarían si al modelo del diseño en panel con efectos simultáneos se le añade una variable exógena constante o no a lo largo del tiempo, según lo visto anteriormente:

Para la exógena constante:

$$X_2 = \alpha_0 + \alpha_1 X_1 + \alpha_2 Y_1 + \alpha_3 Y_2 + \alpha_4 Z + \varepsilon_1 \quad (25)$$

$$Y_2 = \beta_0 + \beta_1 Y_1 + \beta_2 X_1 + \beta_3 X_2 + \beta_4 Z + \varepsilon_2 \quad (26)$$

Para la exógena no constante:

$$X_2 = \alpha_0 + \alpha_1 X_1 + \alpha_2 Y_1 + \alpha_3 Y_2 + \alpha_4 Z_1 + \varepsilon_1 \quad (27)$$

$$Y_2 = \beta_0 + \beta_1 Y_1 + \beta_2 X_1 + \beta_3 X_2 + \beta_4 Z_2 + \varepsilon_2 \quad (28)$$

Estas expresiones presentan la misma problemática que se ha comentado en los modelos no recursivos: no son identificados. Con lo cual será necesario recurrir a la estrategia antes analizada para su reparametrización. Por otro lado, igualmente señalado con anterioridad, la aplicación de la estimación OLS en cada ecuación por separado ofrece soluciones sesgadas. Existe una variedad de soluciones a este respecto, aunque citaremos como exponente suficientemente conoci-

do la solución 2SLS basada en una estimación OLS en dos etapas. En la primera, se obtienen estimaciones de parámetros de las ecuaciones que se plantean mediante el empleo de la definición de las endógenas en función de todas las exógenas. Así, en el segundo caso que acabamos de exponer (exógena no constante), podría establecerse:

$$X_2 = \tau_0 + \tau_1 X_1 + \tau_2 Y_1 + \tau_3 Z_1 + \tau_4 Z_2 \tag{29}$$

$$Y_2 = \tau_0 + \tau_1 X_1 + \tau_2 Y_1 + \tau_3 Z_1 + \tau_4 Z_2 \tag{30}$$

La primera ecuación la representaremos por  $\pi_1$  y a la segunda por  $\pi_2$ . Una vez obtenidas las estimaciones OLS de los parámetros  $\tau_i$ , substituiremos las anteriores ecuaciones en las originales, obteniéndose así:

$$X_2 = a_0 + a_1 X_1 + a_2 Y_1 + a_3 \pi_2 + a_4 Z_1 + \varepsilon_1 \tag{31}$$

$$Y_2 = \beta_0 + \beta_1 Y_1 + \beta_2 X_1 + \beta_3 \pi_1 + \beta_4 Z_2 + \varepsilon_2 \tag{32}$$

Este procedimiento, para la obtención de estimaciones no sesgadas no es,

evidentemente, el único. Como se ha comentado, existe una variedad suficiente de soluciones distintas a esta problemática (3SLS, Ridge Regression, Normalized Ridge Regression, ML...). Sirva simplemente este pequeño ejemplo para la exposición breve de la problemática expuesta.

### MODELO DE ECUACIONES ESTRUCTURALES CON VARIABLES LATENTES: EFECTO DEL ERROR DE MEDIDA EN LOS DISEÑOS EN PANEL

La cuestión del empleo de variables directamente observables, propia del Path Analysis, nos plantea otra problemática relativa a la evaluación del error de medida entre la puntuación obtenida y la puntuación verdadera. Los modelos de ecuaciones estructurales contemplan esa cuestión incluyendo en su definición variables de carácter no observable. Básicamente supone disponer, como mínimo, de dos indicadores para cada uno de los constructos implicados en el diseño, usando para ello el modelo del análisis facto-

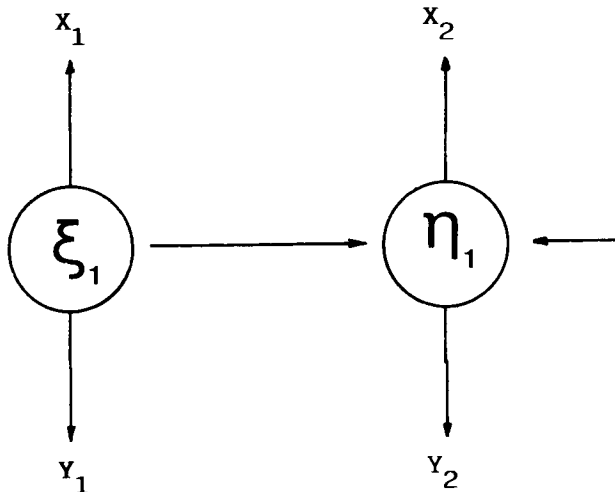


Figura nº 7.- Diagrama inicial del diseño en panel con variables latentes.

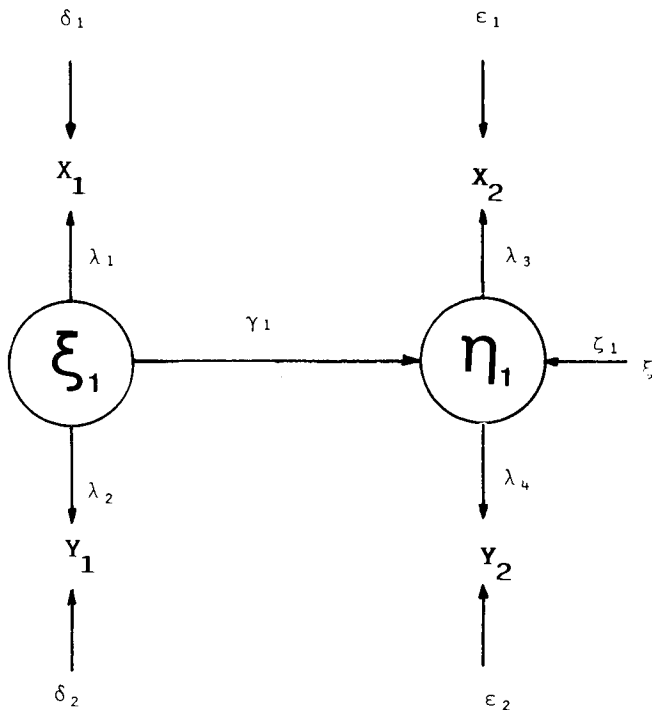


Figura nº 8.- Presentación del Path anterior con la notación Lisrel.

rial para representar cada constructo como una combinación lineal de las variables observables. Inicialmente podría establecerse el siguiente path diagrama:

$\xi_1$  y  $\eta_1$  representan dos constructos que quedan definidos mediante las variables observables que se indican. De cualquier modo, y por cuestiones de notación, el diagrama anterior puede plantearse del siguiente modo (por ceñirnos a la notación clásica de los sistemas de ecuaciones estructurales):

Este diagrama puede traducirse en las siguientes ecuaciones:

$$\eta_1 = \gamma_1 \xi_1 + \zeta_1 \quad (33)$$

$$x_1 = \lambda_1 \xi_1 + \delta_1 \quad (34)$$

$$y_1 = \lambda_2 \xi_1 + \delta_2 \quad (35)$$

$$x_2 = \lambda_3 \xi_2 + \epsilon_1 \quad (36)$$

$$y_2 = \lambda_4 \xi_2 + \epsilon_2 \quad (37)$$

De forma general, los tres modelos implícitos en la configuración anteriormente planteada, responden a las siguientes expresiones:

$$\eta = \beta \eta + \xi + \zeta \quad (38)$$

$$x = \lambda \xi + \delta \quad (39)$$

$$y = \lambda \xi + \epsilon \quad (40)$$

Atendiendo a este planteamiento, simplemente nos limitaremos a plantear dos de los modelos completos correspondientes a los diseños en panel (2W2V).

### 1.- Diseño en panel con efectos simultáneos

La figura número 9 muestra el path diagrama de un diseño en panel con variables latentes, lo cual constituye la extensión estructural del modelo presentado en la figura 4.

Adviértase la diferencia en la notación, pero creemos que con ello nos ajustamos más a lo que se puede consultar en las publicaciones a este respecto. El diagrama correspondiente al Path Analysis, podría establecerse del siguiente modo:

Las ecuaciones estructurales que se desprenden son:

$$\eta_1 = \beta_{21} \eta_2 + \gamma_{11} \xi_1 + \gamma_{12} \xi_2 + \zeta_1 \tag{41}$$

$$\eta_2 = \beta_{12} \eta_1 + \gamma_{21} \xi_1 + \gamma_{22} \xi_2 + \zeta_2 \tag{42}$$

La definición de  $\eta_1$ ,  $\eta_2$ ,  $\xi_1$ , y  $\xi_2$  dependen de las expresiones antes comentadas y que responden al modelo de análisis factorial. La obtención de las estimaciones de los parámetros siguen las disposiciones ya comentadas y en consecuencia no las volveremos a comentar. Simplemente indicar que la posibilidad de utili-

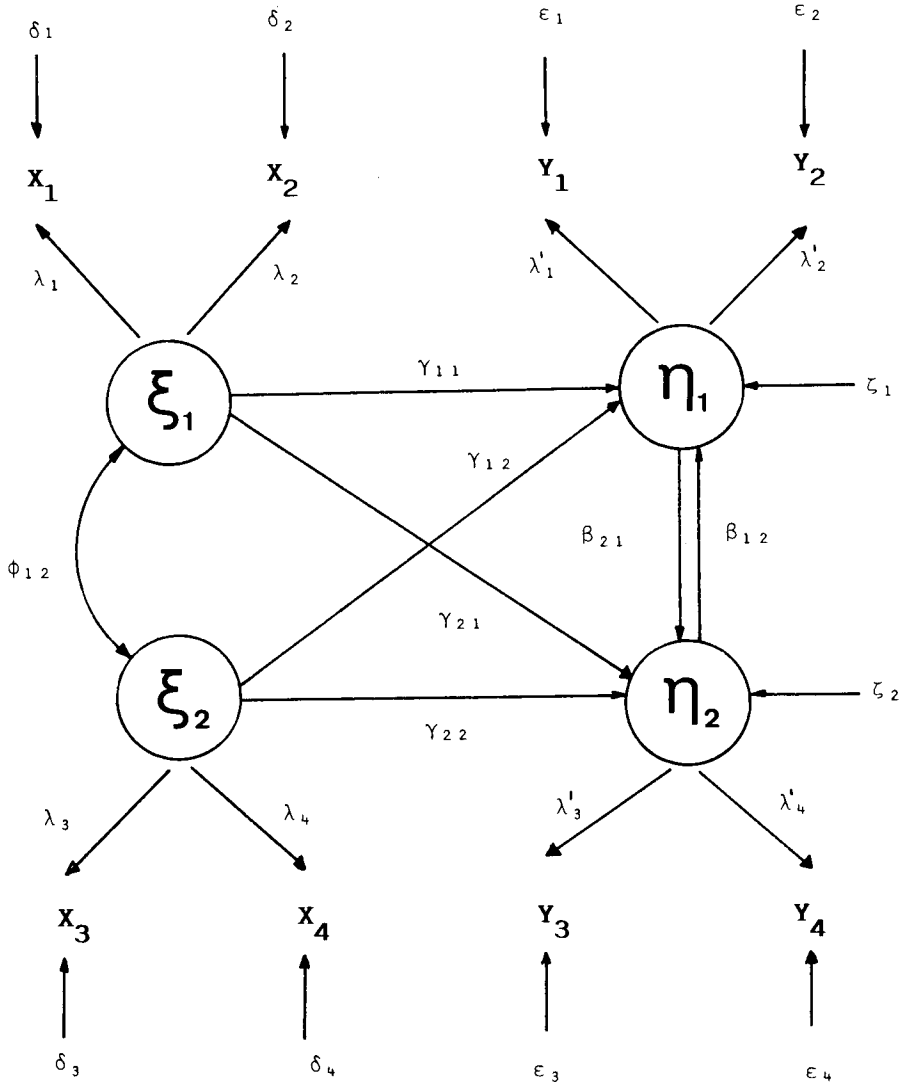


Figura nº 9.- Path diagrama completo del modelo estructural con variables latentes.

zar el programa LISREL VII para la estimación de máxima verosimilitud ofrece una propuesta a considerar, aunque no exenta de complicaciones (Loehlin, 1987).

2.- *Diseño en panel con una exógena constante*

En la figura anterior se muestra el path diagrama correspondiente al modelo planteado con variables latentes y la existencia de una variable exógena ( $\xi_2$ ) para el establecimiento de efectos espurios.

Las ecuaciones correspondientes a este diagrama se expresan como:

$$\eta_1 = \beta_{12} \eta_2 + \beta_{13} \eta_3 + \gamma_{11} \xi_1 + \zeta_1 \tag{43}$$

$$\eta_2 = \gamma_{21} \xi_1 + \gamma_{22} \xi_2 + \zeta_2 \tag{44}$$

$$\eta_3 = \beta_{31} \eta_1 + \beta_{32} \eta_2 + \gamma_{31} \xi_1 + \gamma_{32} \xi_2 + \zeta_3 \tag{45}$$

La definición de las variables latentes responde a la misma configuración ya comentada.

Como se puede observar a partir de la gráfica anterior, la complejidad de es-

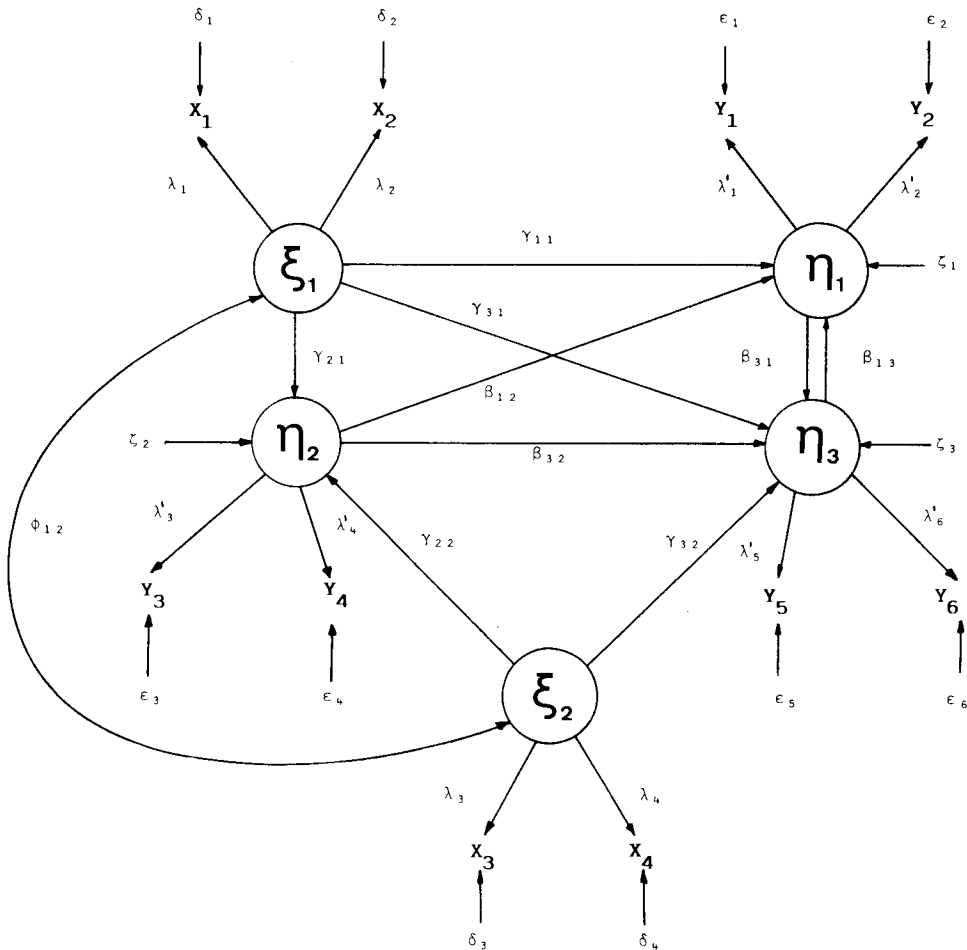


Figura nº 10.- Path Diagrama con variables latentes incluyendo la definición de modelos de medida para todas las variables.

pecificación es evidente. Pero, a la vez, la posibilidad de analizar los distintos efectos entre variables posee una calidad difícilmente conseguible de otro modo. Por citar un ejemplo, el análisis de los valores  $\beta_{12}$  y  $\gamma_{31}$  no puede plantearse de forma individual, puesto que será relevante a este efecto el estudio de las distintas cargas factoriales de las variables latentes que se implican en los dos parámetros citados (Dykstra, 1971).

### CONCLUSIONES

Un último comentario general, relacionado con algunas de las características que se han mencionado a lo largo de las anteriores páginas, parece necesario. A partir de lo expuesto, entendemos que el empleo de los sistemas de ecuaciones estructurales es una alternativa a considerar firmemente para el estudio de diseños longitudinales. Ello no significa que esta práctica no esté exenta de algunas dificultades, como por ejemplo su complejidad con diseños con más de dos medias, que pueden generar ciertas reticencias en su utilización. De forma breve, deberemos

recordar que el coste de la complejidad, por citar el ejemplo anterior, no es desproporcionado si se evalúan las aportaciones de la técnica en cuestión. El establecimiento de múltiples efectos (directos, indirectos, recíprocos, condicionales...) favorecen una mayor flexibilidad en el análisis. La utilización de variables latentes abre el camino a todo aquello que se relaciona con la medida por una parte y con la relación entre medidas por otra. La posibilidad de los efectos recíprocos (no reciprocidad), a pesar de la polémica que suscitan, permiten análisis del efecto temporal mucho más exhaustivos que mediante cualquier otra técnica.

De todos modos, creemos que es necesario recordar que la estrategia de análisis de los modelos estructurales no está exenta de dificultades, no ya de orden aplicado, sino de orden interno. Simplemente recuperar aquí la problemática del ajuste, de su especificación, de la estimación, el efecto del tamaño de muestra, entre otros, para determinar que los sistemas de ecuaciones estructurales merecen una consideración general prudente y delimitar claramente el verdadero alcance que en cada caso puedan presentar los resultados así obtenidos.

### BIBLIOGRAFIA

- Berry, W. D. (1983). *Nonrecursive Causal Models*. Beverly Hills: Sage.
- Campbell, D. T. y Stanley, J. C. (1966). *Experimental and Quasi-Experimental Design for Research*. Chicago: McNally.
- Cook, T. D. y Campbell, D. T. (1979). *Quasi-Experimentation*. Chicago: McNally.
- Dwyer, J. H. (1983). *Statistical Models for the Social and Behavioral Sciences*. New York: Oxford University Press.
- Dykstra, A. (1971). Logic of Causal Analysis: From Experimental to Nonexperimental Designs. En Blalock, H. M. Jr., *Causal Models in Social Sciences*. New York: Aldine-Atherton.
- Guàrdia, J. (1986). *Los modelos de ecuaciones estructurales en el ámbito de la Psicología*. Tesis Doctoral no publicada. Universidad de Barcelona.
- Gujarati, D. N. (1988). *Basic Econometric*. New York: McGraw-Hill Int. Edit.
- Jöreskog, K. G. y Sörbom, D. (1977). Statistical models and methods for analysis of longitudinal data. En D. J. Aigner y A. S. Goldberger (eds.), *Latent Variables in Socio-Economic Models*. Amsterdam: North Holland.
- Kenny, D. A. (1975). Cross-lagged panel co-

- relation: A test for spuriousness. *Psychological Bulletin*, 82, 887-903.
- Loehlin, J. C. (1987). *Latent Variable Models*. Hillsdale: Lawrence Erlbaum Associates.
- Plewis, I. (1985). *Analysing Change*. Chichester: Jhon Wiley and Sons.
- Rogosa, D. (1979). Causal models in longitudinal research: Rationale, Formulation and Interpretation. En J. R. Nesselroade y P. B. Baltes (eds.), *Longitudinal research in the study of Behavior and Development*. New York: Academic Press.
- (1980). A critique of cross-lagged correlation. *Psychological Bulletin*, 88 (2), 245-258.

Dirección de los autores:

Facultat de Psicologia. Divisió de Ciències de la Salut. Universitat de Barcelona. Departament de Metodologia de les Ciències del Comportament. c/. Adolf Florensà, s/n. 08028 Barcelona.