

---

# ANALISIS Y MODELIZACION CAUSAL EN SOCIOLOGIA

Manuel García Ferrando

---

## I. *Introducción*

En este trabajo pretendo adentrarme en, a la vez que clarificar parcialmente, una aparente paradoja. La que se manifiesta al considerar, de un lado, el progresivo descrédito filosófico de la noción de causa, acelerado a partir de finales del siglo pasado y, de otro lado, no sólo la persistencia de esa noción en las Ciencias Sociales, sino el creciente y considerable volumen de publicaciones profesionales dedicadas al análisis causal y a su aplicación en la investigación sociológica.

Esta paradójica situación requiere una clarificación centrada en la respuesta al interrogante siguiente: ¿qué función o utilidad se encuentra en la noción de causa que justifique esa atención creciente que recibe en la investigación sociológica y que parece contra corriente?

A su vez y puesto que se manejan y definen diversas nociones de causalidad, que han sido diseñadas con objetivos y en relación a problemas distintos, la profundización en esa respuesta exige acotar el terreno. Como se verá más adelante, la noción de causa de Simon y su formalización del concepto de orden causal han sido las bases del actual desarrollo del análisis causa y de la modelización causal en el análisis sociológico. Por ello, centraré mi atención

---

preferentemente en el estudio de la técnica del análisis causal, de la modelización causal o del *path analysis* o análisis de camino, que con todos estos nombres se conoce a la tecnología desarrollada a partir de la noción de causa de Simon. Pero antes de adentrarme en el estudio de la técnica del análisis causal, analizaré brevemente el papel que juega la noción de causa en la filosofía de la ciencia y las peculiaridades que presenta la causación social en contraposición a la causación en el mundo natural.

## II. La noción filosófica de causa

Desde que Aristóteles acuñara filosóficamente el término de causa, el análisis filosófico de las causas ha conocido diversos avatares, interesando destacar aquí la crítica epistemológica que realiza David Hume en su conocido *Ensayo sobre el entendimiento humano*. Hume negó que tuviéramos ningún fundamento válido para la atribución de causa y efecto, y nos desafió a encontrar algo más en la naturaleza o en la experiencia que simples covariaciones o a un objeto siguiendo a otro. Aunque Kant replicara convincentemente a Hume en la *Crítica de la razón pura*, la polémica filosófica sobre la noción de causa tomará otros derroteros en tiempos recientes con el desarrollo de la física teórica y de la filosofía positivista<sup>1</sup>.

En efecto, el físico Ernst Mach, junto con otros, propondrá la sustitución de la noción de causa por la noción más general de función<sup>2</sup>. Se trata de un proceso que iniciado por físicos y matemáticos converge con el que protagoniza la filosofía de la ciencia de la mano de Popper, quien en su *Lógica de la investigación científica* propone la reducción de las explicaciones causales a explicaciones por leyes<sup>3</sup>, y de Hanson que en *Patterns of discovery* sostiene la relativización teórica de las nociones causales<sup>4</sup>.

Síntesis de esta situación de crisis filosófica de la noción de causa podría ser la conocida cita de Bertrand Russell, quien llegó a afirmar lo siguiente: «La palabra causa se encuentra tan inextricablemente enlazada con asociaciones engañosas que resulta deseable eliminarla por entero del vocabulario filosófico. Todos los filósofos, de todas las escuelas, imaginan que la causalidad es uno de los axiomas o postulados fundamentales de la ciencia, aunque, cosa singular, la palabra 'causa' nunca aparezca en las ciencias avanzadas. La razón de que la física haya dejado de buscar las causas consiste, en realidad, en que

<sup>1</sup> Para una discusión clara y bien documentada sobre la polémica filosófica de la noción de causa, se puede ver Marx W. WARTOFKY, *Introducción a la Filosofía de la Ciencia*, 2, Madrid, Alianza Universidad, 1976, págs. 379-408.

<sup>2</sup> J. J. KOCKELMANS, *Philosophy of Science*, Nueva York, Free Press, 1968, recoge con precisión esta fase del desarrollo de la filosofía positivista.

<sup>3</sup> Karl R. POPPER, *La Lógica de la Investigación Científica*, Madrid, Tecnos, 1973.

<sup>4</sup> N. R. HANSON, *Patterns of Discovery*, Cambridge, Cambridge University Press, 1968.

no existen tales causas. Pienso que la ley de causalidad, como mucho de lo que vale para los filósofos, es como una *reliquia de edades caducas*, que, al igual que la monarquía, *sobrevive porque erróneamente se supone que no perjudica*.

No cabe duda de que la razón de que la vieja ley de causalidad haya seguido campeando tanto tiempo en los libros de filosofía, es simplemente que los filósofos desconocen la idea de función, y en consecuencia buscan una formulación ilusoriamente simplificada»<sup>5</sup>.

Pues bien, a pesar de las brillantes palabras de Bertrand Russell, en Sociología, y en las Ciencias Sociales en general, existe un hecho irrefutable: el término «causa» ha estado presente siempre en el análisis sociológico. Y del mismo modo que la institución monárquica ha sabido en algunos países, como en España, recrearse en formas nuevas excelentemente adaptadas a la sociedad democrática, así le ha ocurrido al análisis causal en Sociología, que ha adoptado nuevos enfoques epistemológicos y metodológicos mejor adaptados a la lógica de la investigación analítica, basada en buena medida en la investigación por medio de encuestas. Lo que a su vez representa, como se verá más adelante, su flanco débil, que amenaza en convertir al actual análisis causal en Sociología en una tecnología desprovista de toda sustancia teórica explicativa. Y es que muchos de los numerosos artículos profesionales que se publican bajo títulos encabezados por las expresiones «Un análisis causal de...» o «Un análisis de camino de...», se encuentran tan desprovistos de significado teórico y conocimiento empírico relevante, y tan repletos de coeficientes y fórmulas estadísticas que uno no puede por menos preguntarse hasta qué punto es posible reunir bajo la misma etiqueta de trabajo sociológico a tales trabajos, y a los que han seguido más fielmente la tradición científica inaugurada por los padres fundadores de nuestra disciplina.

Pero retornando a la relevancia sociológica de la noción de causa, el hecho evidente es que los seres humanos actúan y piensan dando por supuesto la relación causa-efecto. Como ha mostrado Piaget, desde muy temprano los niños utilizan el lenguaje causal: los primeros usos del por qué se dirigen a inquirir por los motivos de las acciones y la justificación de las reglas. Más tarde, cuando el egocentrismo infantil comienza a declinar, se producen las demandas por explicaciones causales naturalistas. Dado que los seres humanos se encuentran inmersos en el proceso universal e irremediable del cambio, el primer problema de la mente que investiga es preguntarse por el por qué del cambio<sup>6</sup>.

<sup>5</sup> Bertrand RUSSELL, "Sobre la noción de causa", en *Misticismo y Lógica y otros ensayos*, Buenos Aires, Paidós, 1951, págs. 178-179.

<sup>6</sup> Un buen resumen de las breves ideas de Piaget aquí expuestas, y que fueron publicadas originalmente en *Le langage et la pensée chez l'enfant* (1923) y *La causalité physique chez l'enfant* (1927), se encuentra en A. M. BATTRO, *El pensamiento de Jean Piaget*, Buenos Aires, Emecé, 1969.

Como señalara MacIver en su ya clásico trabajo sobre la causación social<sup>7</sup>, el mundo que experimentamos es un mundo de continuidad y de cambio. Percibimos el cambio y, por consiguiente, la continuidad, y viceversa. Es un mundo en el cual el concepto de causación reina sobre nuestra experiencia. Nuestra vida es un proceso señalado por sucesos. Nos pasan cosas y reaccionamos ante lo que nos pasa, y la totalidad de ello es nuestra experiencia. Nadie puede hacer que no hayan sucedido las cosas que han sucedido. Por tanto, nuestra experiencia debe asumir siempre el carácter de un nexo causal. No hay escape de la red de causa y efecto.

Ahora bien, la causalidad del mundo social presenta una importante diferencia en relación a la causalidad en el mundo de la Naturaleza, porque el nexo causal que se atribuye a los cambios en éste difiere en un aspecto importante del nexo causal que se descubre cuando somos nosotros mismos los autores o fuentes del cambio. En efecto, en la sociedad se pueden prever, en alguna medida, los cambios que efectuamos; emprendemos la actividad para llevar a cabo el cambio previsto. La actividad teleológica introduce, pues, un factor de causación que el mundo físico no revela. En el proceso teleológico, el efecto, no ya como realidad, sino como previsión y fin de una acción, opera para efectuarse. Como señala de nuevo MacIver, el efecto, más o menos claro, existe como proyecto antes del proceso físico del cual, como realidad, emerge. Esta relación constituye el factor esencial en el control que el ser vivo o el grupo social ejercen dentro de su medio<sup>8</sup>.

Ahora bien, cuando se investiga el porqué del cambio, el sociólogo está particularmente interesado en el porqué de la unión o de la relación social. Esta, de manera contraria al nexo físico, no existe aparte de los objetivos y motivos de los seres sociales; y de forma diferente también al nexo teleológico, no constituye en sí misma una relación de medios a fines. Así, por ejemplo, cuando se pregunta por qué la tasa de divorcios en las diferentes sociedades está relacionada con el grado de urbanización e industrialización de las mismas, o por qué las tasas de movilidad social en los países industrializados varían de la forma en que lo hacen, el sociólogo responde no con una referencia a los fines que los hombres buscan, sino en el marco de las interacciones y relaciones sociales que se establecen.

Pero no siempre el sociólogo estudia el cambio social de aspectos tan específicos de la estructura social como son las tasas de divorcio o de movilidad social. Los fenómenos sociales más interesantes, con frecuencia, se resisten a ser aislados y definidos con sencillez y facilidad. Los fenómenos sociales suelen estar enredados en la maraña de toda su historia, que hace que dependan de condiciones que nacen dentro de cada orden de la realidad física y socio-histórica<sup>9</sup>. Por qué se desarrolla el capitalismo de la forma en lo que hace

---

<sup>7</sup> Rober M. MAC IVER, *Causación Social*, México, Fondo de Cultura Económica, 1949.

<sup>8</sup> MAC IVER, *op. cit.*, págs. 10-12.

<sup>9</sup> MAC IVER, *op. cit.*, págs. 23-24.

y en unos países concretos y no en otros, es una pregunta que se han hecho, y continúan haciéndose, los sociólogos y cuya respuesta en términos causales requiere una delimitación previa del fenómeno, en el correspondiente marco histórico, de extrema dificultad.

En su teoría del tipo ideal Weber trató de lograr una formulación teórica a los intentos de descripción e interpretación de una situación histórica o de un sistema social. Dado que la comprensión de un fenómeno sociohistórico complejo es inevitablemente parcial y selectiva, un modo de comprensión válido puede basarse en los modos típicos de la conducta, de las normas generalmente aceptadas de la acción, que, al mismo tiempo, son atribuibles a los actos de los individuos y se reflejan en la estructura institucional del sistema. Los individuos tienden a actuar de acuerdo con tales normas, y es esta tendencia la que da al sistema su forma y cualidad típicas. De esta tendencia se derivan «leyes» de un orden o de una etapa histórica determinada. Como afirma Weber, «tales leyes son determinadas probabilidades típicas, confirmadas por la observación de que, dadas determinadas situaciones de hecho, transcurran con la forma esperada ciertas acciones sociales que son comprensibles por sus motivos típicos y por el sentido típico mentado por los sujetos de la acción»<sup>10</sup>.

Sin entrar en estos momentos en las complejidades y ambigüedades de la teoría del tipo ideal de Weber<sup>11</sup>, lo que me interesa destacar aquí es que al asignar un papel decisivo a uno o varios factores como determinantes del cambio del fenómeno social que se pretende explicar, del conjunto de factores que inciden en el mismo, Weber señala un camino enormemente fructífero para el análisis causal del cambio social, al concebirlo como un orden comprensible de las cosas.

De lo que se trata, según Weber, es de captar la realidad social como algo coherente, como una unidad de elementos, en toda época histórica, aun admitiendo que se presentan numerosas contradicciones y excepciones. Se persigue el tipo, la comprensión de muchas particularidades dentro de una unidad, aunque cada tipo sea un caso único.

Inspirándose en el método interpretativo weberiano, Mac Iver formuló más recientemente sus propuestas para un análisis de la causación social<sup>12</sup>. Según este autor, al concebir una situación social como un sistema de fuerzas afines que la sostienen, surge la interrogación de la causación cuando una situación cesa de conformarse a nuestra expectativa, cuando no revela más, como antes, sus actividades y probabilidades típicas. Entonces cabe presumir

<sup>10</sup> MAX WEBER, *Ensayos sobre metodología sociológica*, Buenos Aires, Amorrortu, 1973, págs. 262 y ss.

<sup>11</sup> Weber fue ambivalente, cuando no ambiguo, en muchos de sus estudios metodológicos. Por eso son de mayor importancia sus estudios histórico-sociológicos concretos para evaluar la metodología weberiana. De opinión parecida es Irving ZEITLIN, *Ideología y Teoría Sociológica*, Buenos Aires, Amorrortu, 1970, pág. 138.

<sup>12</sup> MAC IVER, *op. cit.*, págs. 213 y ss.

que la estructura dominante de fuerzas que la sostenía ha sufrido un serio cambio. Y entonces se busca lo que Mac Iver denomina factor precipitante, es decir, la fuerza que se haya introducido. Así, para Weber, la ética protestante fue el disolvente de las restricciones tradicionales a la empresa de negocios y a la adquisición monetaria, y por ello, una causa primordial en el nacimiento del capitalismo.

Ahora bien, resulta más difícil relacionar el precipitante con una actitud que con un suceso, de ahí que la teoría weberiana sobre los orígenes del capitalismo ha sido cuestionada desde muchas perspectivas<sup>13</sup>. El propio Robert Mac Iver sostiene que para poder admitir la explicación de Weber es preciso antes refutar la explicación que sostiene que el nacimiento de la ética protestante —con su individualismo, su ascetismo mundano, y su noción de la gestión económica— fue la expresión, en la esfera religiosa, del correspondiente cambio total en las actitudes sociales, cambio que era interdependiente de un orden socioeconómico cambiante<sup>14</sup>.

Pero en la contraargumentación de Mac Iver a la tesis weberiana, subyace la propia debilidad de la aproximación interpretativa del análisis de la causación social, ya que difícilmente se puede demostrar con la suficiente convicción que un determinado factor causal juega realmente un papel decisivo como precipitante. Es decir, el problema consiste en cómo justificar en la explicación del cambio social, la selección entre la compleja totalidad de la estructura social de ciertos elementos como factores causales decisivos del cambio. Según Mac Iver, la tarea de la investigación causal debe fundamentarse en el método comparativo inspirado en el diseño experimental, aun admitiendo que prácticamente nunca podrá el sociólogo emprender auténticos experimentos sociales como hace el científico natural cuando experimenta con la materia física. Para ello, y después de especificar y deslindar al máximo el por qué social que se desea responder e interpretar, el sociólogo debe identificar la situación o tipo de situación en la cual se produce el fenómeno, en contraste con una situación o tipo de situación comparable pero en la que no se produce el fenómeno. Si  $X$  es la diferencia específica y se encuentra dentro de la situación o coyuntura  $C$ , se procederá a la consideración de  $C_1$ , la situación o coyuntura comparable a la que falta  $X$ . Entonces se establecerá un marco de referencia que abarque tanto a  $C$  como a  $C_1$ . La investigación preliminar sobre las causas se dirige al establecimiento de una diferencia específica  $X$  y de aquí a la determinación más precisa de las situaciones, más íntimamente comparables,  $C$  y  $C_1$ , que seamos capaces de descubrir.

Dado el carácter único e irreplicable de los sucesos o hechos históricos, y dada la imposibilidad de realizar experimentos sociales, el sociólogo debe utilizar, según Mac Iver, el procedimiento universal de la causación con sumo

<sup>13</sup> Recordemos aquí por la contundencia de sus argumentos contra las tesis weberianas la obra del sueco Kurt SAMUELSON, *Religion and Economic Action*, Estocolmo, Scandinavian University Books, 1957.

<sup>14</sup> MAC IVER, *ibidem*.

cuidado ensayando siempre que sea posible la reconstrucción social de los fenómenos sociales, de tal suerte que se puedan interpretar los aspectos teleológicos de los mismos. Ahora bien, en último término, las fórmulas propuestas por Weber, Mac Iver y aquellos otros sociólogos que se apoyan en una interpretación de los fenómenos sociohistóricos, conceden una decisiva importancia a las propias capacidades personales de los científicos sociales, a su sagacidad, erudición, formación intelectual e intuición, a la par que son completamente ambiguas por lo que se refiere al establecimiento de criterios específicos y contrastables que aseguren la necesaria fiabilidad al trabajo científico realizado.

Tal imprecisión va a contracorriente del carácter analítico y cuantitativo que ha ido adquiriendo la ciencia en general, y la sociología en particular, en el transcurso del siglo xx. De ahí que las enseñanzas de Max Weber, y más modestamente de Mac Iver, se utilicen por los actuales sociólogos más como grandes marcos generales de inspiración científica, que como modos de trabajo científico concretos en los que basar la rutina del quehacer cotidiano. Es por ello que el análisis de la causación social tal como se practica hoy en día, descansa más en las fórmulas operativas de los Lazarsfeld, Simon, Blalock o Duncan, basadas como están en una concepción correlacional y estadística, y por tanto cuantificable, de la causalidad. Veamos esto con más detalle.

### III. *Correlación y causación: El modelo de Simon*

Los trabajos de Paul Lazarsfeld sobre la lógica del análisis de encuestas y particularmente el análisis de las relaciones estadísticas en la investigación empírica<sup>15</sup>, estimularon el interés de Herbert Simon por el estudio de la causalidad y de las correlaciones espurias. A su vez, el seminal trabajo de Simon, publicado en 1952 en la revista de la *Asociación Americana de Filosofía*, titulado «On the Definition of the Causal Relation»<sup>16</sup>, seguido del publicado dos años más tarde en el *Journal of the American Statistical Association*, con el título de «Spurious correlation: A causal interpretation»<sup>17</sup>, estimularon el trabajo, entre otros, de econométricos y sociólogos, cristalizando todo ello en lo que se ha convertido en una tecnología ampliamente utilizada en la actualidad en la investigación social empírica bajo la etiqueta de «análisis causal».

Bien es cierto que este tipo de análisis causal tiene unos orígenes anteriores a los aquí señalados, ya que Sewall Wright, padre del *path analysis* o análisis

<sup>15</sup> La obra lazarsfeldiana teórica y metodológica sobre estos temas es amplia y diversa. Pero quizá sea la obra coeditada con ROSENBERG, *The Language of Social Research* (Nueva York, Free Press, 1955) la que recoja más sistemáticamente sus contribuciones fundamentales al análisis sociológico empírico.

<sup>16</sup> Herbert A. SIMON, "On the Definition of the Causal Relation", *The Journal of Philosophy*, 49, 1952, págs. 517-528.

<sup>17</sup> Herbert A. SIMON, "Spurious Correlation: A Causal Interpretation", *Journal of the American Statistical Association*, 49, 1954, págs. 467-479.

de camino, investigando en el campo de la genética publicó, en 1921, un artículo titulado «Correlation and Causation», en el que estableció las bases conceptuales y estadísticas del análisis de ecuaciones estructurales utilizando coeficientes de correlación y regresión<sup>18</sup>. Pero sólo será con el trabajo de Lazarsfeld y Simon, y más adelante con los modelos causales de Blalock, que el análisis correlacional de la causalidad se convierta en una de las metodologías más utilizadas en la investigación social empírica.

El trabajo de Simon es particularmente interesante porque está ligado directamente al análisis causal en sociología y a la modelización en econometría, disciplina ésta que se ha convertido en el paradigma a seguir por aquellos sociólogos interesados en la metodología «dura» de la ordenación causal articulada con el análisis de camino o *path analysis*.

Veamos ahora con cierto detalle los fundamentos de las aportaciones de Simon y Blalock al análisis causal. El concepto de ordenación causal que sostiene Simon<sup>19</sup> no pretende solucionar el «problema de Hume» ni mucho menos contradecir la afirmación humeana de que todo lo que se puede observar son simples covariaciones. En lugar de enfrascarse en un planteamiento ontológico de la noción de causa, esto es, en una noción basada en la conexión «necesaria» de los hechos, Simon se esfuerza por seguir un camino operativo que le permita abordar el análisis causal desde un concepto de fácil manejo por todos los investigadores sociales, como es el de correlación. Con ello, qué duda cabe, no pretende demostrar, en el sentido absoluto del término, la existencia de la causación, pero en cambio ofrece al investigador empírico un instrumento eficaz para enfrentarse desde una concepción en buena medida intuitiva y de sentido común de la causación, con el estudio relacional de las variables sociológicas.

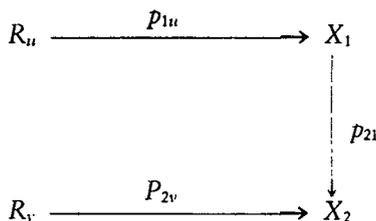
Simon sitúa el estudio de la noción de causalidad en el contexto teórico de la explicación científica. De este modo pretende delimitar claramente la noción de causalidad por un lado y de correlación por otro. La correlación entre dos variables viene establecida por los datos empíricos, mientras que la relación causal es postulada por la teoría, esto es, la correlación sirve para predecir y la causalidad, además, para explicar, y lo que trata Simon es de establecer las condiciones bajo las cuales una correlación entre dos variables diferente de cero, ofrezca la evidencia suficiente para inferir la existencia de una relación causal entre dichas variables. Para introducir su noción de causalidad, Simon estudia en primer lugar la situación bivariable, afirmando que «la correlación es una prueba de causación en el caso de dos variables si podemos establecer los supuestos de precedencia en el tiempo y de la ausencia de correlación entre los términos residuales».

---

<sup>18</sup> Sewall WRIGHT, "Correlation and Causation", *Journal of Agricultural Research*, 20, 1921, págs. 557-585.

<sup>19</sup> El argumento que a continuación se desarrolla es un resumen de los conceptos básicos que presenta SIMON en sus dos artículos anteriores.

De una forma esquemática, la situación que estudia Simon se puede representar en un diagrama de flechas y mediante ecuaciones estructurales del siguiente modo:



en donde  $X_1$  y  $X_2$  son las variables medidas en el modelo, siendo  $X_1$  una *variable exógena*, ya que no se encuentra influida por otras variables del modelo, mientras que  $X_2$  es una *variable endógena*, ya que se encuentra afectada por la *variable*  $X_1$ . Los  $R_i$  son términos de error o residuales, y representan aquellos factores no medidos que inciden sobre las variables endógenas. Los coeficientes  $p$  representan el impacto de una variable sobre otra.

El diagrama de flechas así representado se puede convertir en un sistema de ecuaciones que reflejen las uniones representadas en el diagrama. A cada variable endógena le corresponde una ecuación que incluye aquellas variables que directamente inciden sobre ella. Pues bien, el diagrama anterior se puede representar mediante el sencillo sistema de ecuaciones:

$$X_1 = P_{1u} R_u$$

$$X_2 = P_{21} X_1 + P_{2v} R_v$$

Los supuestos avanzados por Simon en su noción de causalidad demandan, en primer lugar, que para que exista una relación causal entre  $X_1$  y  $X_2$ , sea posible especificar *a priori* la dirección de la relación. En el diagrama se observa que  $X_1$  influye sobre  $X_2$  y no al revés, y el coeficiente  $P_{21}$  representa el impacto causal de  $X_1$  en  $X_2$ . El segundo supuesto exige que los términos residuales asociados con  $X_1$  y  $X_2$ , esto es,  $R_u$  y  $R_v$ , no estén correlacionados entre sí, lo que equivale a decir que no existe ninguna variable que incida simultáneamente sobre  $X_1$  y  $X_2$ . Satisfechos estos supuestos o restricciones, Simon demuestra que la estimación de  $P_{21}$ , esto es, del coeficiente que mide el impacto causal de  $X_1$  sobre  $X_2$ , es simplemente la correlación  $r_{12}$  entre  $X_1$  y  $X_2$ .

Antes de seguir adelante con la discusión del modelo de Simon al introducir una tercera variable, conviene que nos detengamos unos momentos en la consideración de la propia noción de causalidad en la que se basan Simon y el gran número de autores que han continuado su línea de trabajo. Pues bien, para tales autores, es preciso especificar tres condiciones para poder inferir la existencia de una relación causal entre dos variables  $X$  e  $Y$ . La primera

condición establece tal como se ha visto anteriormente, que debe existir una variación concomitante o covariación entre  $X$  e  $Y$ , mientras que la segunda condición exige una asimetría temporal u ordenación en el tiempo de ambas variables. Estas dos condiciones no son excesivamente restrictivas, ya que con frecuencia se puede medir la covariación y observar o imponer una secuencia temporal entre dos variables. Sin embargo, hay una tercera condición que hace más problemática la noción de causalidad, ya que requiere la eliminación de otros posibles factores causales que puedan incidir en la relación observada entre  $X$  e  $Y$ . Dicho de otra forma, la tercera condición establece que para que exista una relación causal entre  $X$  e  $Y$ , la covariación entre ambas variables no debe desaparecer cuando se controlan los efectos de otras variables que pudieran ser causalmente previas a  $X$  e  $Y$ .

Ahora bien, si para cualquier tipo de fenómeno resulta difícil aislar otros posibles factores causales de la relación causal entre  $X$  e  $Y$  que se postula, el problema se agrava para el caso de las ciencias sociales por las especiales dificultades metodológicas que presenta la delimitación clara y precisa del cambio social. Pues tal como han señalado tantos sociólogos, desde Weber hasta Mac Iver, cualquier cambio social que se trate de explicar se halla enredado en la maraña de toda su historia. Depende de condiciones que nacen de los diversos órdenes de la realidad y a menudo, los mismos fenómenos difícilmente se pueden delimitar de manera unívoca, ya que presentan un aspecto diferente a los diversos observadores, de acuerdo con sus propios intereses y posicionamientos sociales. Por todo ello, los fenómenos sociales con frecuencia no ofrecen límites claros que indiquen dónde termina uno y comienza el otro.

Pero todas estas limitaciones no impiden, aunque sí dificultan, la selección de variables a introducir en los modelos explicatorios. Una prudente utilización de los ya amplios conocimientos teóricos y empíricos acumulados sobre gran parte de los fenómenos sociales, son la mejor garantía de que se incluirán sólo aquellas variables relevantes causalmente relacionadas, aunque el investigador social ha de estar preparado para identificar aquellas variables que puedan incidir en la relación original, para incorporarlas (y medirlas) en el modelo. Por eso resulta inevitable formular algunos supuestos restrictivos *a priori* antes de proceder al contraste empírico de cualquier diseño de investigación, siendo inevitable postular que los efectos de otras variables no incluidas en el modelo son despreciables. Más adelante se insistirá en los supuestos *a priori* que acompañan a las propiedades de los términos residuales o de error en los modelos estructurales, pero ahora vamos a continuar con la elaboración del modelo de Simon para el caso de tres variables.

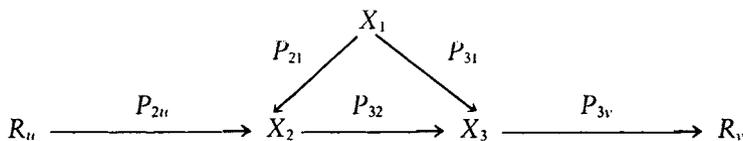
Una vez calculado el coeficiente de correlación entre dos variables, si se sospecha que la correlación observada tiene un origen espurio, se puede introducir una tercera variable  $Z$  que conjeturamos pueda dar cuenta de la correlación observada. Si al calcular el coeficiente de correlación parcial entre  $X$  e  $Y$  manteniendo constante la tercera variable introducida, el resultado se aproxima

a cero, mientras que no lo era la correlación original, caben dos tipos de interpretación: que la tercera variable sea interviniente, esto es, que el efecto causal de  $X$  e  $Y$  opere a través de  $Z$  o que la correlación entre  $X$  e  $Y$  resulte del efecto causal de  $Z$  en ambas variables, con lo que la relación original es espuria.

Algunas veces, el propio contexto de la investigación permite seleccionar una de las dos interpretaciones. Así, Simon, utilizando dos ejemplos adaptados del clásico trabajo de Zeisel *Dígalo con números*<sup>20</sup>, presenta su argumento del siguiente modo. El consumo de dulces se encuentra negativa y altamente relacionado con el *status* matrimonial. Pero también existe una correlación negativa y alta entre el consumo de dulces y la edad, mientras que la relación entre la edad y el *status* matrimonial es también alta pero positiva. Al mantener la edad constante, sin embargo, la correlación entre *status* matrimonial y consumo de dulces se hace prácticamente cero. Ahora caben hacer dos interpretaciones, la edad es una variable interviniente entre el consumo de dulces y el *status* matrimonial, o bien la correlación entre estas dos variables es espuria, ya que en realidad se trata de un efecto conjunto de la edad. Nuestro sentido común nos indica que la última interpretación es la correcta, ya que el consumo de dulces depende de la edad y no del *status* matrimonial.

Supongamos ahora que encontramos una correlación alta y positiva entre el porcentaje de mujeres que trabajan y que están casadas y el índice de absentismo laboral. Pero si introducimos en el análisis correlacional el número de horas dedicadas al trabajo doméstico, la correlación original desaparece. De nuevo, nuestro sentido común nos señala la dirección correcta de la causación al considerar las variables horas de trabajo doméstico como una variable interviniente entre la población laboral femenina casada y absentismo; esto es, que el matrimonio de las mujeres tiende a aumentar la cantidad de trabajo doméstico que realizan, lo que a su vez provoca un mayor absentismo laboral.

Pero no siempre se puede recurrir al sentido común para dilucidar si una correlación es genuina, y por tanto implica causalidad, o por el contrario se trata de una correlación espuria que no implica relación causal alguna. Aunque en último término es siempre el marco teórico el que nos permite elegir entre las dos alternativas, el establecimiento de un diagrama causal y de un sistema de ecuaciones estructurales, permite fortalecer nuestro razonamiento lógico. El modelo de Simon nos ayuda a tomar correctamente tales decisiones. Para el caso de tres variables se puede escribir el siguiente diagrama:



<sup>20</sup> Hans ZEISEL, *Dígalo con Números*, México, Fondo de Cultura Económica, 1974 (e.o. 1947).

Suponemos que la correlación entre  $X_2$  y  $X_3$  es diferente de cero y queremos saber si  $P_{32}$  es también diferente de cero, lo que significaría que  $X_2$  incide causalmente sobre  $X_3$ . Así, pues, lo que analiza Simon son las condiciones bajo las que desaparece  $P_{32}$ .

A partir del anterior diagrama, se puede escribir el siguiente sistema de ecuaciones estructurales:

$$\begin{aligned} X_2 &= P_{21} X_1 + P_{2u} R_u \\ X_3 &= P_{31} X_1 + P_{32} X_2 + P_{3v} R_v \end{aligned}$$

En su trabajo original, Simon emplea un procedimiento un tanto complejo de multiplicación de pares de ecuaciones para estimar  $P_{32}$  y a continuación examina las condiciones bajo los que  $P_{32}$  vale cero.

Una forma más rápida y sencilla de estimación del coeficiente  $P_{32}$  sugerida por Stokes es la del procedimiento de la variable instrumental, que consiste, en resumen, en operar en cada ecuación estructural de forma que se obtenga un sistema de  $m$  ecuaciones con  $n$  incógnitas —los coeficientes  $P_{ij}$ —, siendo  $m \geq n^{21}$ . Ahora bien, tanto se siga un procedimiento u otro, es preciso satisfacer los supuestos apriorísticos de que las variables residuales no están correlacionadas entre sí, y tampoco lo están con las variables explicatorias. Se trata de restricciones que no son verificables estadísticamente y que hay que aceptar como válidas para poder operar con el modelo.

Se puede demostrar que la solución para el coeficiente  $P_{32}$ , siguiendo el método de Simon o el procedimiento de la variable instrumental de Stokes es que

$$P_{32} = \frac{r_{23} - r_{12} r_{13}}{1 - r_{12}^2} = \beta_{32 : 1}$$

El rasgo más importante de esta solución es que el valor del coeficiente  $P_{32}$  es equivalente al peso beta  $\beta_{32:1}$ , que es el coeficiente estandarizado (o tipificado) de regresión. También resulta de interés señalar que el numerador de un coeficiente estandarizado de regresión es el mismo que el numerador del correspondiente coeficiente de correlación parcial, con lo que Simon observa que el numerador del coeficiente  $P_{32}$  es igual al numerador del coeficiente de correlación parcial  $r_{23:1}$ . Por tanto, si beta tiende a cero, lo mismo le ocurrirá al correspondiente coeficiente parcial  $r$ , con lo que la comprobación del posible carácter espurio de una correlación se reduce a la predicción de un valor cero o aproximadamente cero de ciertos coeficientes de correlación parcial. Simbólicamente se tiene que para una secuencia lineal tal como  $X_1 \rightarrow X_2 \rightarrow X_3$ ,

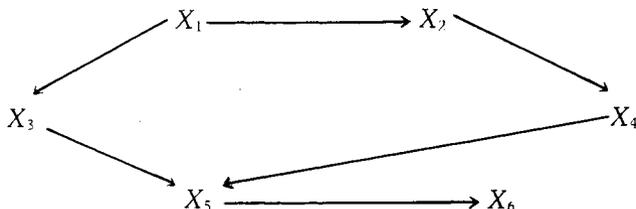
<sup>21</sup> D. E. STOKES, "Compound paths: an expository note", *American Journal of Political Science*, 18, 1974, págs. 191-214.

la predicción es que  $r_{13.2}=0$ , mientras que para que se produzca una secuencia espuria tal como  $X_1 \begin{matrix} \nearrow X_2 \\ \searrow X_3 \end{matrix}$  la predicción ha de ser  $r_{23.1}=0$ .

Posteriormente al trabajo original de Simon, Blalock ha sugerido que en lugar de los coeficientes de correlación, que ofrecen a veces predicciones imprecisas, se utilicen coeficientes de regresión que son más fiables<sup>22</sup>. Sin embargo, en la práctica, las diferencias entre utilizar un tipo de coeficiente u otro no son significativas.

Pero la aportación más interesante de Blalock ha consistido en apoyarse en el modelo de Simon para desarrollar una técnica que permite comprobar la existencia o no de uniones entre variables en modelos recursivos de cualquier tamaño. La secuencia que se sigue en la técnica de Blalock comienza por construir un modelo en forma de diagrama de flechas, para después observar dónde se han omitido las uniones entre pares de variables y a continuación generar predicciones sobre ciertos coeficientes de correlación parcial referentes a dichos pares de variables, que debieran aproximarse al valor cero. Entonces se calcula el valor real de la correlación parcial, se compara con el valor predicho, y en base a la comparación se decide si se debe o no omitir la unión<sup>23</sup>.

El procedimiento lógico ideado por Blalock quizá se pueda seguir mejor si lo desarrollamos a través de un ejemplo práctico de su uso. En un estudio ya clásico con este tipo de técnica sobre la conducta electoral del público norteamericano, Golberg trató de encontrar un modelo teórico que explicase dicho comportamiento<sup>24</sup>. El primer modelo que sometió Goldberg al contraste de los datos empíricos fue el siguiente, tras el examen previo de los conocimientos teóricos y empíricos acumulados sobre el tema:



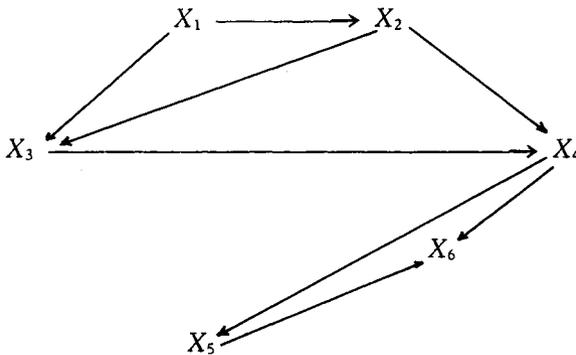
<sup>22</sup> Hubert M. BLALOCK, "Correlated Independent variables: the problem of multicollinearity", *American Journal of Sociology*, 42, 1963, págs. 233-237.

<sup>23</sup> BLALOCK es un autor prolífico y sus aportaciones al análisis causal se encuentran dispersas en no menos de diez publicaciones diferentes. Como más características se pueden citar aquí su libro *Causal Inferences in Nonexperimental Research*, Chapel Hill, University of North Carolina Press, 1964, y el artículo "Theory Building and Causal Inference", págs. 155-198, en H. M. BLALOCK y A. B. BLALOCK (eds.), *Methodology in Social Research*, Nueva York, McGraw-Hill, 1968.

<sup>24</sup> A. GOLDBERGER, "Discerning a causal pattern among data on voting behavior", págs. 33-48, en H. M. BLALOCK, Jr. (ed.), *Causal Models in the Social Sciences*, Chicago, Aldine-Atherton, 1966.

en donde la variable  $X_1$  representa las características sociológicas del padre del votante,  $X_2$  es el partido al que está afiliado el padre del votante,  $X_3$  son las características sociológicas del votante,  $X_4$  es el partido al que está afiliado el votante,  $X_5$  son las actitudes del votante frente a los partidos y  $X_6$  es el voto del votante, o variable que se trata de explicar.

Las ecuaciones predictivas teóricas que se deducen del modelo requieren que sean cero los siguientes coeficientes de correlación parcial:  $r_{41.23}$ ;  $r_{61.2345}$ ;  $r_{32.1}$ ;  $r_{52.134}$ ;  $r_{62.1345}$ ;  $r_{43.12}$ ;  $r_{63.12}$ ;  $r_{63.1245}$ , y  $r_{64.1235}$ . Calculados los valores reales, Goldberg encontró que los coeficientes  $r_{32.1}$ ,  $r_{43.12}$  y  $r_{64.1235}$  eran mayores de 0,1, lo que condujo a revisar el modelo y a incluir las uniones entre  $X_2$ - $X_3$ ,  $X_3$ - $X_4$  y  $X_4$ - $X_6$ :



De este modo se obtiene el modelo que según Goldberg minimiza las discrepancias entre los valores predichos y los observados.

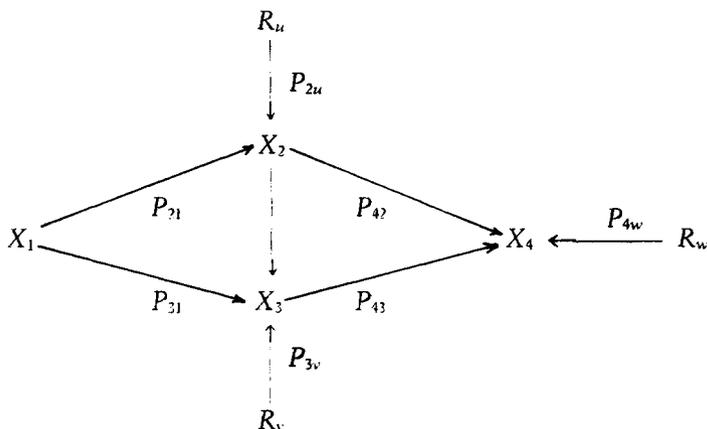
Vemos, pues, que la técnica de Blalock lo único que ofrece es información acerca de si se debe incluir o no una unión entre dos variables, aunque poco dice sobre la fuerza de la unión entre ambas. A cambio de este tipo de resultado es preciso formular unos supuestos previos muy restrictivos que han de satisfacer las variables introducidas en el modelo. En efecto, al basarse la técnica de Blalock en el modelo básico de regresión, su utilización adecuada requiere que se satisfagan los supuestos del análisis de regresión, esto es, que los datos se encuentren medidos a nivel de intervalo, que los términos residuales no se encuentren correlacionados con las variables independientes, que el modelo sea lineal y que no se hayan producido errores aleatorios en la medición de las variables. Además de estos supuestos correspondientes al análisis de regresión, la técnica de Blalock exige que el sistema de ecuaciones lineales sea recursivo, esto es, que no contenga relaciones de doble dirección. En suma, se trata de satisfacer supuestos altamente restrictivos que raras veces se cumplen en puridad en la investigación empírica en sociología, a cambio de obtener unos resultados en el mejor de los casos poco brillantes.

Es por ello que a pesar de la popularidad de la técnica de Blalock en la modelización causal en sociología, se haya impuesto, sin embargo, en los últimos años otra técnica que aunque impone las mismas restricciones previas que la técnica de Simon-Blalock, ofrece a cambio una información resultante más rica para el investigador.

#### IV. *El análisis de camino*

El análisis de camino se ocupa básicamente de estimar la magnitud de las uniones entre variables, utilizando dichas estimaciones para suministrar información acerca de los procesos causales subyacentes. La forma más simple de obtener los coeficientes de camino es mediante el empleo de técnicas comunes de regresión, siempre que se cumplan los supuestos que exigen tales técnicas, sobre todo la restricción de que la variable residual en una ecuación estructural no se encuentre correlacionada con las variables explicatorias de la ecuación<sup>25</sup>.

A guisa de ejemplo y sin salirnos de la sociología política, tomemos el modelo desarrollado por Miller y Stokes<sup>26</sup> para predecir la conducta de los representantes políticos elegidos en los correspondientes distritos electorales:



<sup>25</sup> Son muy numerosos los trabajos aparecidos en la última década sobre el análisis de camino. Para una presentación esquemática de esta técnica se puede consultar M. GARCÍA FERRANDO, *Socioestadística*, 2.ª edición, Madrid, CIS, 1984, que contiene además en el capítulo XIV una bibliografía más ampliada.

Sobre el problema de los modelos estructurales con variables no observadas, se puede consultar el trabajo de Francisco ALVIRA y Marina PEÑA, "Path Analysis: Modelos de ecuaciones estructurales y variables no observadas", *REIS*, 3, 1978, págs. 187-208.

<sup>26</sup> W. E. MILLER y D. E. STOKES, "Constituency influence in Congress", págs. 351-372, en A. CAMPBELL et al., *Elections and the Political Order*, Nueva York, 1966.

en donde  $X_4$  es la conducta del representante elegido,  $X_1$  son las actitudes de los electores,  $X_2$  la percepción que los representantes tienen de las actitudes de los electores, y  $X_3$  son las actitudes de los representantes.

Sin entrar en los detalles técnicos del análisis del camino, digamos que las ecuaciones estructurales que se deducen del modelo son las siguientes:

$$\begin{aligned} X_2 &= P_{21} X_1 + P_{2u} R_u \\ X_3 &= P_{31} X_1 + P_{32} X_2 + P_{3v} R_v \\ X_4 &= P_{42} X_2 + P_{43} X_3 + P_{4w} R_w \end{aligned}$$

El procedimiento de estimación de los principales coeficientes de camino consiste sencillamente en regresar cada variable endógena en aquellas variables que directamente inciden sobre ella, mientras que los coeficientes de camino residuales, esto es,  $P_{2u}$ ,  $P_{3v}$  y  $P_{4w}$ , también se extraen siguiendo técnicas de regresión mediante la fórmula  $\sqrt{1 - R^2}$  en donde  $R^2$  es el cuadrado del correspondiente coeficiente de correlación múltiple.

La utilidad real del análisis de camino radica en el uso que se puede hacer de las estimaciones de los coeficientes de camino. De esta manera es posible medir los efectos directos e indirectos que una variable tiene sobre otra. Así, en el ejemplo anterior, las actitudes de los electores tienen además del efecto directo, un efecto indirecto en la conducta de los representantes vía su impacto en las actitudes del representante y de su percepción de las actitudes de los electores. Entonces se puede comparar la magnitud de los efectos directos e indirectos que permitirán identificar los mecanismos causales operativos.

Además, el análisis de camino permite descomponer la correlación entre cualquier par de variables en una suma de coeficientes de camino simples y compuestos, siendo algunos de estos coeficientes compuestos efectos indirectos significativos y otros no. La descomposición de la correlación se puede hacer de diversas formas, aunque la más operativa e intuitiva de todas, consiste en seguir las reglas que Wright estableciera en un seminal artículo publicado en 1934<sup>27</sup> y que tuvo que esperar al desarrollo y difusión de los ordenadores de alta velocidad, para que las ideas y conceptos vertidas en él pudieran ser utilizadas más fácilmente por los actuales investigadores sociales y no sociales.

En el modelo de Miller y Stokes se puede descomponer, por ejemplo, la correlación entre  $X_1$  y  $X_3$ , diferenciando entre el efecto directo entre ambas variables, lo que vendría representado por el coeficiente  $P_{31}$ . Como también se puede ir de  $X_1$  a  $X_3$  vía  $X_2$ , el correspondiente coeficiente de camino sería

<sup>27</sup> Sewal WRIGHT, "The Method of Path Coefficients", *Annals of Mathematical Statistics*, 5, 1934, págs. 161-215.

$P_{21} \cdot P_{32}$ . De este modo, la descomposición de  $r_{13}$  es igual a la suma de los coeficientes de camino simple y compuesto, esto es

$$r_{13} = \underbrace{P_{31}}_{\text{efecto directo}} + \underbrace{P_{21} \cdot P_{32}}_{\text{efecto indirecto}}$$

La descomposición del coeficiente de correlación suministra además una forma de comprobación de cuán adecuado es el modelo si algunas uniones se han omitido en el modelo inicial, ya que en tal caso el coeficiente de camino correspondiente valdrá cero. De este modo, el *test* o contraste del modelo consiste sencillamente en comprobar si el coeficiente omitido es en efecto cero, determinando la magnitud de la unión emitida mediante la estimación del modelo que incluya dicha unión.

Vemos, pues, que el análisis recursivo de camino ofrece idéntica información que la técnica Simon-Blalock más cierta información adicional, y todo ello con las mismas restricciones o supuestos *a priori*. En resumen, el análisis de camino permite examinar los procesos causales que subyacen a las relaciones observadas y también permite estimar la importancia relativa de caminos alternativos de influencia. El tipo de contraste del modelo que se puede realizar mediante el análisis de camino estimula una mejor aproximación causal en la búsqueda de explicaciones de los fenómenos investigados.

Tales ventajas del análisis de camino lo han convertido en una de las técnicas más utilizadas en la investigación social y empírica. Económetras y sociómetras han empleado profusamente dicha técnica, en sus intentos por descubrir y contrastar pretendidas relaciones causales entre los fenómenos sociales y económicos<sup>28</sup>. Pero vistas las cosas con la perspectiva que ofrece el paso del tiempo, el balance actual de la utilización del análisis de camino en la investigación sociológica no parece en mi opinión excesivamente brillante. Miller y Stokes se han preocupado recientemente de estudiar el uso que se ha hecho del análisis de camino en los trabajos de investigación publicados en las tres principales revistas profesionales de sociología americanas<sup>29</sup>. Pues bien, en el período estudiado, que va de 1965 a 1973, 48 artículos publicados en tales revistas han utilizado el análisis de camino en 174 modelos elaborados con un número diverso de variables (sólo dos en el modelo más sencillo y 13 en el modelo más complejo).

Pues bien, con el fin de evaluar el uso realizado del modelo de camino en los referidos artículos, Miller y Stokes calcularon el valor del término resi-

<sup>28</sup> Para una detallada presentación de la técnica de Simon-Blalock y de sus aplicaciones en las ciencias sociales, se puede ver Herbert B. ASHER, *Causal Modeling*, Beverly Hill, Sage Publications, 1976.

<sup>29</sup> M. K. MILLER y C. S. STOKES, "Path analysis in sociological research", *Rural Sociology*, 40, 1975, págs. 390-394.

dual ( $1 - R^2$ ), que representa el porcentaje de varianza explicado por las variables que no *intervienen* en el modelo. Sus resultados fueron sorprendentes. En los modelos que habían utilizado alrededor de seis variables, el valor medio del término residual fue de 0.79. Un residuo de esta magnitud indica que por término medio se ha explicado menos del 40 por 100 de la varianza. Otros resultados del examen del empleo del análisis de camino fueron que mientras que el 20 por 100 de los modelos tenían términos residuales menores de 0.70, otro 25 por 100 tenía residuos incluso superiores a 0.90. Dicho de otra forma, un artículo de cada cuatro explica menos del 20 por 100 de la varianza con una media de cinco variables. Además, casi un 11 por 100 de tales artículos no pudieron explicar más del 10 por 100 de la varianza total, e incluso otro 10 por 100 de los artículos no incluyeron el término residual ni ningún otro coeficiente que hubiera permitido calcular la varianza no explicada. Todos estos datos, como afirman Miller y Stokes, despiertan un fundado escepticismo en los modelos cuantitativos basados en el análisis de camino tal como vienen siendo utilizados en la investigación sociológica.

Bien es cierto que el porcentaje de varianza explicada e inexplorada por las variables introducidas en el modelo no debe utilizarse como la prueba definitiva de la adecuación del modelo al que se le ha aplicado la técnica del análisis de camino. Otis Dudley Duncan dejó bien claro, en un trabajo publicado hace ya casi diez años<sup>30</sup>, que el objeto de la investigación científica en general y de la técnica del análisis de camino en particular, no puede consistir sólo en maximizar la varianza explicada. Más bien considera Duncan que el logro de un sistema sólido de ecuaciones algebraicas es secundario en relación a la obtención de una representación (o modelo) apropiada de la estructura del problema. Esto es, que lo importante no son las soluciones algebraicas, ni siquiera el preciso ajuste algebraico de los sistemas de ecuaciones, sino el logro de una interpretación racional y sustantiva de los datos.

#### V. *Las insuficiencias teóricas de la modelización causal*

Sin embargo, las cosas no parecen marchar tal como hubieran deseado los fundadores de esta técnica de análisis. Porque desde Wright hasta Blalock, pasando por Simon, siempre se ha destacado que la contribución de la modelización causal y de los sistemas de ecuaciones estructurales del análisis de camino, no consiste tanto en la racionalización de los cálculos de la varianza explicada o en la evaluación de la «relativa importancia» de las variables, como en lograr explicar la formulación de los supuestos que deben de preceder a tales cálculos si se quiere que ofrezcan resultados inteligibles. Como ya

<sup>30</sup> Otis DUDLEY DUNCAN, *Introduction to Structural Equation Models*, Nueva York, Academic Press, 1975. En el capítulo 2 presenta Duncan una precisa discusión de las diferencias y analogías entre los conceptos de correlación y causalidad.

señalara Wright en un trabajo publicado en 1931, el poder real del análisis de camino no radica en el tipo de cálculos de regresión que realiza, sino en su utilización para estudiar las propiedades de sistemas más complejos que los que permite una simple regresión lineal. Como señala de nuevo Duncan, en problemas en los que este tipo de sistemas es el modelo apropiado, el cálculo de las varianzas explicadas o es irrelevante o es un objetivo secundario.

Pero la creciente aparición en las revistas profesionales sociológicas de artículos abstractos que contienen el análisis de camino como la técnica básica para el análisis de los datos, lo que por otro lado está haciendo que tales revistas resulten cada vez más aburridas y obtusas, parece indicar que muchos sociólogos han emprendido el estudio de los modelos de ecuaciones estructurales con la esperanza de adquirir una técnica que se pueda aplicar mecánicamente a una serie de datos numéricos, en la creencia de que el resultado se convierte automáticamente en una «pieza de investigación científica». Este deseo de encontrar la clave de la «sociología instantánea», como señala agudamente Otis Dudley Duncan en su básico libro anteriormente citado, y esa absurda suposición de que cualquier conjunto de datos, torturados según el ritual estadístico prescrito en los programas de cálculo para ordenador, van a producir descubrimientos científicos interesantes, están anclados en la falacia de la inducción, es decir, en la suposición de que el conocimiento científico se basa en reglas de inducción. Pero, afortunadamente, en el campo de la filosofía de la ciencia ya nadie, o casi nadie, defiende en la actualidad la postura induccionista, que se encuentra en total desprestigio<sup>31</sup>.

La historia reciente de las investigaciones sociológicas que se publican en libros y revistas, y que utilizan modelos de ecuaciones estructurales, suministra, pues, otro episodio de empleo de la última moda metodológica en la búsqueda de la «sociología instantánea». Pero la investigación sociológica, cuando produce resultados sustantivos y relevantes sobre el funcionamiento de los fenómenos sociales, está basada en modelos inspirados en teorías sociológicas creativas, sustantivas y profundas.

Y si no se dispone de modelos tales, la más abstracta y elegante de las formalizaciones apenas servirá para otra cosa que conseguir que sus autores superen el listón de los estándares establecidos por los editores de las revistas profesionales, y logren publicar trabajos que aportan bien poco al avance del conocimiento científico y sociológico.

Tal es la situación que ha alcanzado el área del «análisis causal» en las ciencias sociales, que el propio Herbert Simon ha hecho oír recientemente su autorizada voz para tratar de introducir algo de orden en el exuberante y

---

<sup>31</sup> Como prueba de la escasa importancia del tema de la inducción, baste señalar que en las discusiones que se producen desde las dos grandes posturas que dominan el actual panorama de la filosofía de la ciencia, esto es, la filosofía analítica y la teoría crítica, apenas si se producen referencias sobre el problema de la inducción.

tumultuoso crecimiento de la bibliografía dedicada al análisis de camino y a la identificación y estimación de ecuaciones estructurales<sup>32</sup>.

Enlazando con las ideas que en su día expusiera convincentemente Paul Lazarsfeld acerca de la importancia del conocimiento previo para el establecimiento de atribuciones causales, Simon realiza una aportación al libro homenaje al propio Lazarsfeld coeditado por sus amigos y colaboradores Merton, Coleman y Rossi, tratando de clarificar el significado de la ordenación causal. Frente a los autores que se muestran preocupados casi obsesivamente por el ajuste algebraico del sistema de ecuaciones estructurales, Simon señala que el problema de los coeficientes de las ecuaciones de regresión no es un problema estadístico que desaparecería, aunque fuéramos capaces de medir con total exactitud y de observar todas las variables implicadas en las relaciones postuladas. Incluso en tales circunstancias, es preciso formular supuestos *a priori* acerca de los mecanismos que subyacen en el comportamiento de las variables, si deseamos estimar tales coeficientes.

Pero aquí surge aparentemente una contradicción de apariencia circular. Se afirma que hay que realizar restricciones *a priori* con el fin de construir modelos identificados de los fenómenos que se trata de explicar. Ahora bien, las características especiales de las teorías científicas, y que las diferencia de las teológicas y de las metafísicas, es que deben basarse en hechos y poder ser contrastadas por la evidencia empírica. En tal caso, pues, ¿qué queda de los fundamentos empíricos de la ciencia si se exigen restricciones *a priori* en el meollo de la construcción teórica?

Este tema ha provocado una amplia discusión en el campo de la metodología causal, discusión que ha sido zanjada convincentemente por Newell y Simon<sup>33</sup>. Ambos autores argumentan que los supuestos apriorísticos de la modelización causal utilizados en la identificación de las ecuaciones lineales, no son previos a toda experiencia: son tan sólo previos a los datos particulares que tratamos de analizar. Representan leyes de estructura cualitativa, derivadas inductivamente de la experiencia científica acumulada.

Este punto conecta, una vez más, con el argumento que sostiene la imposibilidad de reducir el análisis causal a una tecnología de uso automático. Porque aunque, en efecto, se trate de una tecnología, como lo es el análisis factorial o las técnicas de contraste de hipótesis, su correcta utilización no puede reducirse a seguir ordenadamente los pasos que señalan los libros-recetarios que desgraciadamente abundan dentro de la bibliografía metodológica y estadística en sociología. Una comprensión de sus bases conceptuales y un conocimiento exacto de los límites de su validez y aplicabilidad, son las únicas garantías de que el análisis causal pueda servir realmente para el avance del

<sup>32</sup> Herbert A. SIMON, "The meaning of Causal Ordering", págs. 65-81, en R. K. MERTON, J. COLEMAN y P. H. ROSSI, *Qualitative and Quantitative Social Research*, Nueva York, Free Press, 1979.

<sup>33</sup> A. NEWELL y H. A. SIMON, "Computer Science as Empirical Inquiry: Symbols and Search", *Communications of the ACM*, 19, 1976, págs. 113-126.

conocimiento sociológico. Lo que sólo se puede lograr situando los modelos causales en el correspondiente marco teórico, e interpretando los resultados obtenidos a partir de las hipótesis y supuestos formulados previamente. La obtención de sistemas de ecuaciones estructurales bien ajustadas algebraicamente y la explicación de unos porcentajes más o menos elevados de varianza, son resultados completamente secundarios frente al diseño y ajuste de modelos, que permitan interpretar las relaciones entre variables en términos de los mecanismos causales subyacentes que producen tales relaciones. Sin un conocimiento de tales mecanismos, no podemos predecir cómo van a covariar las variables cuando el sistema estudiado experimenta algún tipo de cambio.

Como señalara el propio Simon, la determinación de los mecanismos causales que controlan los fenómenos, es una tarea inductiva. Ningún conjunto finito de datos puede determinar un modelo causal particular y excluir todos los demás. Bien es cierto que esto no nos debe preocupar en demasía, pues ya sabemos desde Popper que las teorías nunca se pueden establecer irrevocablemente, sino tan sólo se pueden refutar por la evidencia.

Por eso, el análisis causal en sociología no debiera seguir, en mi opinión, el sendero previamente trazado por la Econometría. El diseño de modelos causales multivariantes recursivos y recíprocos, que dan lugar a no menos complejos sistemas de ecuaciones estructurales, en cuya solución algebraica mediante programas estandarizados de ordenador parecen diluir lo mejor de sus esfuerzos sus autores, no puede ser un fin en sí mismo para una ciencia social. Más bien, la preocupación del investigador social debe ser la de diseñar modelos que puedan ser interpretados sustantivamente a la luz del conocimiento científico acumulado y de las hipótesis teóricas conjeturadas *a priori*.

Como ya señalara Mac Iver en los años cuarenta, el problema causal en las ciencias sociales no es de *cuánto*, sino de *cómo*. Y debemos tratar de responder de acuerdo con este criterio. Yo diría que los esfuerzos que se dediquen a conocer el cuánto sólo se justifican en tanto que contribuyan a una mejor comprensión del cómo de la atribución causal. Los resultados estadísticos no ofrecen una prueba positiva en la verificación de las hipótesis causales, pero en cambio sí proporcionan formulaciones exactas de las correlaciones y covariaciones de las variables, lo que a su vez facilita saber dónde y cómo debemos buscar las conexiones causales<sup>34</sup>.

En realidad, lo que está ocurriendo con la modelización causal en sociología cabe entenderlo dentro del evidente y preocupante desfase que se ha producido entre el desarrollo de la construcción de teorías, elaboración y medición de indicadores y variables, y la consiguiente obtención de datos por un lado, y el desarrollo de las técnicas de análisis de datos por otro. Estas últimas han alcanzado niveles de sofisticación muy elevados gracias al desarrollo de

<sup>34</sup> Para una panorámica general de la actual polémica metodológica entre las perspectivas cuantitativa y cualitativa en Sociología, se puede ver el trabajo de Francisco ALVIRA, "Perspectiva cualitativa-perspectiva cuantitativa en la metodología sociológica", *REIS*, 22, 1983, págs. 53-76.

los programas estadísticos computarizados que permiten resolver fácilmente los sistemas algebraicos más complejos. Esto es lo que ha permitido que econométricas, psicómetras y sociómetras elaboren modelos en los que se introduce un número considerable de variables, de los que obtienen resultados muy significativos estadísticamente. Pero desgraciadamente, un resultado con una significación estadística del 0.001 puede que no resulte en absoluto significativo desde el punto de vista del sentido común, y mucho menos desde el punto de vista teórico, que es lo que ocurre con frecuencia.

Porque el desarrollo teórico y la construcción de indicadores válidos y fiables de los conceptos claves en la sociología, se encuentran si no estancados sí al menos muy retrasados con respecto a las técnicas analíticas. Así, por ejemplo, todavía no existe una teoría suficientemente precisa que explique el comportamiento político de los ciudadanos. Los modelos causales elaborados hasta ahora por los politólogos que se han inspirado en Simon, no consiguen explicar más allá del 30 ó 40 por 100 de la variación del comportamiento político<sup>35</sup>. Ahora bien, para obtener resultados más sustantivos, la solución no debe seguir el camino de resolver nuevos sistemas de ecuaciones estructurales, sino el de desarrollar modelos teóricos más sustantivos y elaborar indicadores fiables de nuevas variables sociológicas y psicosociológicas que vengán a enriquecer la limitada provisión de variables explicativas que se vienen utilizando, sobre todo en la investigación basada en técnicas de encuesta.

Por todo ello entiendo que paralelamente a los avances en la tecnología de la modelización causal, es preciso dedicar mayores esfuerzos teóricos a la elaboración y medición de conceptos que permitan una mayor y mejor interpretación sociocultural de los fenómenos sociales. Sólo en este doble plano estadístico y teórico-interpretativo cabe esperar nuevas y significativas interpretaciones causales en sociología, evitándose así que la Sociología caiga en los mismos errores cometidos en la Econometría, en donde el formalismo prevalece demasiado a menudo sobre la sustancia, y el logro de la significación estadística oscurece la búsqueda de la significación teórica.

---

<sup>35</sup> El referente obligado de la tendencia modelizadora en la sociología y en la ciencia política es la obra de Herbert A. SIMON, *Models of Man, Social and Rational Human Behavior in a Social Setting*, Nueva York, Wiley, 1957. Para una panorámica general de los problemas metodológicos en el campo de la sociología política, se puede ver M. GARCÍA FERRANDO, "Problemas metodológicos y Técnicos de Investigación en la Sociología Política", *REIS*, 7, 1979, págs. 35-61.