

los procesos de selección sugiriendo, a la vez, estrategias imperfectas pero útiles que pueden darnos cierto control sobre el problema que plantea el estudio. Al fin y al cabo, nuestro objetivo es lograr un diseño de investigación que seleccione en función de las variables explicativas presentes en nuestra teoría y que permita el cambio de la variable dependiente. Sin embargo, al encaminarnos hacia esa meta, puede que sea útil utilizar diseños que tengan en cuenta los valores observados de la variable dependiente; pero al investigador que lo haga hay que aconsejarle suma cautela. Nuestro objetivo primordial es obtener más información relevante para contrastar nuestra teoría, sin por ello introducir un sesgo que ponga en peligro la calidad de las inferencias.

5. Entender qué debe evitarse

En el capítulo 4 analizamos cómo había que elaborar un estudio mediante un diseño de investigación preciso en el que los procedimientos de selección de observaciones hicieran posibles inferencias causales válidas. Sin embargo, realizar bien esta tarea es necesario pero no suficiente para extraer tales inferencias, ya que hay posteriores errores de análisis que pueden destruir el trabajo que antes se ha hecho bien. En este capítulo vamos a estudiar de qué manera, una vez seleccionadas las observaciones para el análisis, es posible comprender las causas de la ineficiencia y del sesgo y reducirlas a proporciones manejables. Posteriormente, nos ocuparemos de cómo podemos controlar la investigación con el fin de enfrentarnos adecuadamente a estos problemas.

Para analizar la falta de eficiencia y el sesgo vamos a recordar los criterios de evaluación de inferencias que presentamos en el apartado 7 del capítulo 2 y en el 4 del capítulo 3. Si tenemos un diseño de investigación preciso, tenemos que centrarnos en los dos problemas clave que se van a estudiar en este capítulo: *sesgo* e *ineficiencia*. Para entender estos conceptos resulta útil imaginarse las inferencias como estimaciones de puntos concretos flanqueados por los extremos de un intervalo. Por ejemplo, podríamos suponer que una persona tiene cuarenta años, dos arriba, dos abajo. Esta cifra es nuestro mejor cálculo (la estimación), que el intervalo que hay entre treinta y ocho y cuarenta y dos sitúa en su centro, con una estimación de nuestra incertidumbre (la amplitud del intervalo). Queremos elegir el intervalo de manera que la verdadera edad esté dentro de él durante mucho

tiempo. *La falta de sesgo tiene que ver con el hecho de centrar un intervalo cerca de la estimación correcta, mientras que la eficiencia se relaciona con el estrechamiento de un intervalo que se ha centrado correctamente.*

Estas definiciones de falta de sesgo y de eficiencia son siempre aplicables, independientemente de que se quieran hacer inferencias descriptivas—como en el ejemplo de la edad— o causales. Si, por ejemplo, queremos hacer una estimación del efecto que tienen los estudios en la renta (los ingresos monetarios que se perciben por cada año de estudio), tendremos un punto de estimación del efecto que estará flanqueado por un intervalo que refleja nuestra incertidumbre respecto a la cantidad exacta. Necesitaremos un intervalo que sea lo más estrecho posible (para una mayor eficiencia) y cuyo centro esté cerca de la estimación correcta (para evitar sesgos); también será necesario que el cálculo aproximado del ancho del intervalo represente sinceramente nuestra incertidumbre.

En este capítulo nos centramos en las causas del sesgo y de la ineficiencia, desde ese primer momento del estudio en el que pretendemos mejorar la calidad de la información hasta cuando llegamos a extraer inferencias causales. En el apartado 1 analizamos los errores de medida que pueden sesgar nuestros resultados y hacerlos también menos eficientes. Posteriormente, en el apartado 2, pasamos a considerar el sesgo que puede aparecer en nuestras inferencias causales cuando hemos eliminado del análisis variables explicativas que se deberían haber incluido. En el apartado 3 abordamos el problema contrario: controlar las variables irrelevantes que reducen la eficiencia de nuestro análisis. A continuación, estudiaremos el problema que se plantea cuando nuestra variable «dependiente» afecta a las «explicativas». Este problema se conoce con el nombre de «endogeneidad» y se aborda en el apartado 4. Para terminar, en los apartados 5 y 6 analizaremos, respectivamente, la adjudicación aleatoria de valores en las variables explicativas, así como diversos métodos de control no experimental.

1. Error de medida

Una vez que hemos seleccionado nuestras observaciones, tenemos que medir los valores de las variables que nos interesan. Como en las ciencias sociales todas las observaciones y medidas son imprecisas, nos encontraremos enseguida con el problema del error de medida.

En la investigación social la mayoría de los análisis intentan calcular el nivel de error y reducirlo lo más posible. Los estudios cuantitativos disponen de medidas más precisas (numéricas), pero no necesariamente más certeras. La fiabilidad (el que diversas medidas del mismo fenómeno tengan el mismo resultado) a veces se obtiene sacrificando la validez (que las medidas reflejen lo que el observador trata de calibrar). Los cualitativistas intentan obtener medidas certeras que, sin embargo, suelen ser un poco menos precisas.

En ciertos rasgos esenciales, la observación cualitativa y la medición cuantitativa son muy parecidas. Es cierto que en la primera las categorías se suelen etiquetar con palabras, mientras que en la segunda se asignan valores numéricos a las categorías y medidas; sin embargo, los investigadores de una y otra tendencia utilizan criterios nominales, ordinales y de intervalo. Con los de tipo nominal, se conforman grupos de observaciones en los que no se presupone que las categorías tengan un determinado orden. Las que son de relevancia pueden basarse en el ordenamiento legal o institucional; por ejemplo, en la política comparada, pueden ser de interés las pautas de los sistemas presidenciales, parlamentarios o autoritarios en diversos países. Las categorías ordinales dividen los fenómenos siguiendo algún tipo de orden. Por ejemplo, un investigador cualitativo puede clasificar los países en tres o cuatro categorías, teniendo en cuenta su grado de industrialización o el tamaño de sus fuerzas armadas. Para terminar, la medición a través de intervalos utiliza variables continuas, como se hace en los estudios de flujos comerciales internacionales.

Las diferencias que hay entre la medición cuantitativa y la cualitativa afectan a la presentación de los datos, no a la posición teórica de dicha medición. Los cualitativistas utilizan en sus medidas palabras como «más» o «menos», «mayor» o «menor», «fuerte» o «débil», mientras que los cuantitativistas usan números.

Por ejemplo, la mayoría de los investigadores cualitativos del campo de las relaciones internacionales es muy consciente de que el «número de muertos en combate» no tiene por qué ser un buen indicador de la importancia que tienen las guerras para las tendencias posteriores de la política mundial. En la teoría del equilibrio de poderes se considera que, para calibrar la estabilidad, el concepto teórico relevante no es la gravedad de la guerra sino el cambio que se produce en los principales actores «a consecuencia» de ella (véanse Gulick, 1967, y Waltz, 1979, p. 162). Sin embargo, al querer evitar la falta de validez, el cualitativista corre el riesgo de que los errores de medida le hagan no ser fiable. ¿Cómo podemos saber lo que se produce «a consecuencia» de algo si no definimos precisamente esa expresión? De hecho, ésta parece implicar que la valoración se hace en función de un resultado sistémico (que sesgará las estimaciones posteriores de la relación, llevándolas hacia la hipótesis).

No hay ninguna fórmula que pueda determinar cómo se logra el equilibrio entre el uso de indicadores cuantitativos, que quizá no reflejen de forma válida los conceptos subyacentes que nos interesan, y el de valoraciones cualitativas siempre imprecisas y sujetas a sesgos inconscientes. Sin embargo, los investigadores de una y otra tendencia tienen que hacer una estimación de la incertidumbre en sus inferencias. Los cuantitativistas han de presentar, junto a sus medidas numéricas, cuáles son los errores típicos, mientras que los cualitativistas tienen que hacer una estimación de la incertidumbre, redactando cuidadosamente lo que piensan de sus observaciones.

La diferencia entre las mediciones de unos y otros estriba en su forma de presentar ideas que, en esencia, son las mismas.

Las mediciones cualitativas y cuantitativas también se parecen en otro sentido. En ambas, las categorías suelen ser mecanismos creados por el investigador y no vienen «dadas» por la naturaleza. La clasificación de las naciones en regímenes democráticos y autocráticos o parlamentarios y presidenciales depende de elaboraciones intelectuales, al igual que ordenarlas según criterios como el de mayor o menor industrialización.

Es evidente que no existen respuestas con validez universal: todas las mediciones dependen del problema que el investigador pretenda comprender. Lo mejor es que los criterios que forman las categorías estén lo más cerca posible de las primeras ideas teóricas y empíricas del investigador; sin embargo, esta misma pretensión recalca todavía más el hecho de que dichas categorías son mecanismos al servicio de los propósitos del investigador. El número de regímenes parlamentarios en los que la proporcionalidad es el principal sistema de representación depende de cómo se clasifiquen esos «regímenes parlamentarios» y de lo que se entienda por sistema de representación proporcional. En el campo de las relaciones internacionales es posible investigar los flujos monetarios registrados, pero el uso de una medida continua depende de qué transacciones se haya decidido computar, de las reglas que determinan lo que es una transacción y de cómo se definen las fronteras nacionales. Del mismo modo, en unas elecciones al Congreso estadounidense, la proporción de voto demócrata en una circunscripción depende de que, en la clasificación que haga el investigador para su estudio, se presuponga que las etiquetas «demócrata» y «republicano» tienen el mismo significado en las 435 circunscripciones.

Incluso los criterios de categorización que hemos utilizado en este apartado para hacer mediciones (nominales, ordinales o de intervalo) dependen de para qué fines teóricos se utilicen éstas. Por ejemplo, podría parecer evidente que la etnia es un prototipo de variable nominal que, en los Estados Unidos, podría codificarse como negro, blanco, latino, indio y asiático. Sin embargo, existe una gran diversidad dentro de estos grupos nominales, en función de la intensidad con que los miembros de cada grupo se identifican con él. Por lo tanto, podríamos categorizar a los grupos étnicos según una escala ordinal, basándonos, por ejemplo, en la proporción de miembros de cada uno de esos grupos que está muy identificado con aquel al que pertenece. Quizá nos interese también el tamaño de estos grupos: por lo tanto, podríamos utilizar la etnia como criterio de medida para determinar ciertos intervalos. Lo principal es utilizar la medida más apropiada para nuestros fines teóricos.

Los problemas de medición más frecuentes aparecen cuando se mide sin hacer referencia explícita a ninguna estructura teórica. Por ejemplo, a veces los investigadores toman una variable que es continua de forma natural y que puede medirse bien, como la edad, y establecen categorías de juventud,

mediana edad y vejez. Para ciertos fines, esta clasificación puede ser suficiente, pero, como presentación teórica de la edad de una persona, el procedimiento resulta innecesariamente impreciso. El error de agrupación que se producirá será considerable y hay que evitarlo. El hecho de librarse de este tipo de error nos remite al principio de que no hay que descartar datos sin que haya razón para ello.

Sin embargo, podemos cometer el error contrario: asignar valores continuos y numéricos de intervalo a variables que por naturaleza son discontinuas. En general, la medición mediante intervalos no es mejor que la ordinal o nominal. Por ejemplo, en una encuesta se puede preguntar sobre la filiación religiosa y sobre el grado de compromiso en este sentido. La intensidad del segundo indicador podría medirse —si las preguntas están bien planteadas— mediante una variable ordinal, incluso utilizando intervalos, dependiendo de cuál sea la naturaleza del instrumento de medición. Sin embargo, no tendría mucho sentido hacer una lista numérica con las religiones a las que pertenece cada individuo. En este caso, probablemente no exista una variable ordinal o continua, y este procedimiento generaría un error de medida.

Elegir entre categorías nominales, por un lado, u ordinales o de intervalo, por otro, puede implicar un intercambio entre riqueza descriptiva y facilidad a la hora de comparar. Pensemos, por ejemplo, en las normas de voto de las organizaciones internacionales. El ordenamiento institucional que las dicta es importante porque refleja una concepción de la soberanía de los estados y porque tiene consecuencias sobre la clase de resoluciones que puede aprobarse, los recursos que se destinan a la organización y las expectativas de que se cumplan las resoluciones.

Un conjunto de categorías nominales podría distinguir entre aquellos sistemas en los que un único miembro tiene la facultad de vetar cualquier resolución (como ocurría en el Consejo de la Sociedad de Naciones, según el artículo 15 de su Acta de Constitución), aquellos en los que sólo pueden hacerlo ciertos miembros (como en el Consejo de Seguridad de las Naciones Unidas), en los que se decide en función de algún tipo de mayoría cualificada (como en las decisiones que afectan al mercado interno de la Unión Europea) y en los que es una mayoría simple la que toma las decisiones (como ocurre en la mayoría de las votaciones de la Asamblea General de las Naciones Unidas). Cada sistema puede generar sus propias dinámicas de negociación, y, si lo que pretendemos es estudiar cómo funciona uno de ellos en particular (aquel en el que un miembro puede ejercer el veto), es esencial que definamos nuestras categorías para no incluir en el análisis otros tipos de sistemas. Las categorías nominales serían apropiadas para este propósito.

Sin embargo, también podríamos ver estas categorías de forma ordinal, yendo de las más restrictivas (en las que se requiere unanimidad) a las menos (en las que se precisa mayoría simple). Esta categorización sería nece-

saría si tuviéramos que comprobar proposiciones teóricas que tuvieran que ver con la relación que existe entre lo restrictivas que son las normas de voto y las pautas de negociación o con el carácter distributivo de los resultados típicos. No obstante, al menos dos de nuestras categorías —las que contemplan el veto de ciertos miembros y la decisión por medio de mayorías cualificadas— son bastante confusas, ya que suponen la existencia de un abanico de diversos arreglos. La primera categoría permite que un solo miembro ejerza un veto total —práctica que raya en lo dictatorial— y también que veten todos los miembros, a excepción de los pocos que apenas tienen peso; en la segunda se incluye la reglamentación de la Unión Europea, que impide que dos estados constituyan una minoría de bloqueo en decisiones que afectan al mercado interno. La fórmula que utiliza el Fondo Monetario Internacional es, desde el punto nominal, un ejemplo de decisión por mayoría cualificada, pero, en la práctica, concede a los Estados Unidos y, desde hace poco, a la Unión Europea, que actúa como un bloque, la capacidad de constituir este tipo de minoría de bloqueo. De ahí que pueda pertenecer a cualquiera de las dos categorías.

En consecuencia, podríamos intentar dar un paso adelante y hacer una medida intermedia que se basara en la proporción de estados (o de recursos: según el producto nacional bruto, las aportaciones a la organización o la población) que se precisa para aprobar resoluciones, clasificando las organizaciones internacionales según lo restrictivos que sean sus procedimientos de voto.

Sin embargo, con diferentes baremos —por ejemplo, si se calibran los recursos en función de la población o del producto nacional bruto— se obtendrán resultados diferentes, por lo que las ventajas que pueda tener la precisión de tales medidas se verán limitadas por la arbitrariedad de los criterios de medición o por la complejidad de los datos agregados. Cada categoría tiene ventajas e inconvenientes: los fines del investigador deben determinar qué opción se toma.

En los dos siguientes subapartados vamos a analizar las consecuencias específicas del error de medida para la investigación cualitativa y llegaremos a conclusiones que pueden parecer sorprendentes. Casi nadie discutiría que un error *sistemático* de medida, como el que supone exagerar de manera constante el valor de ciertas unidades, causa sesgo y también incoherencia, ya que dicho sesgo no puede desaparecer si se incorporan más observaciones cargadas de errores. Sin embargo, si analizamos el problema con más atención, veremos que sólo ciertos tipos de error de medida sistemático sesgan nuestras inferencias causales. Además, las consecuencias del error de medida *no sistemático* están menos claras. Vamos a analizar este último tipo de error en dos partes: en la variable dependiente y, después, en la explicativa. Tal como vamos a demostrar, el error en la variable dependiente produce ineficiencias que pueden generar resultados incorrectos en cualquier situación y dificultar el rastreo de pruebas constantes de la exis-

tencia de efectos sistemáticos. Dicho de otro modo, el error de medida no sistemático en la variable dependiente no produce sesgo, pero puede aumentar considerablemente la ineficacia. Más interesante resulta la presencia de este tipo de error en la variable causal clave, pues siempre sesga las inferencias de forma predecible. Comprender la naturaleza de estos sesgos nos ayudará a atenuarlos o quizá a evitarlos.

1.1 Error de medida sistemático

En este apartado nos ocupamos de las consecuencias del error de medida *sistemático*, que consiste en exagerar de manera constante el valor de ciertos tipos de unidades y que, a veces, puede sesgar y hacer incoherente la estimación de los efectos causales. Nuestro cometido es determinar qué tipos de sesgo produce cada error de medida sistemático. Tanto en la investigación cuantitativa como en la cualitativa, este error puede derivarse del hecho de que los investigadores tomen decisiones que hagan que los datos favorezcan la constatación de sus expectativas. En el trabajo cuantitativo, puede que utilicen datos sesgados porque son las únicas series numéricas disponibles, mientras que en el cualitativo el error de medida sistemático puede provenir de evaluaciones subjetivas realizadas por investigadores que ya tengan hipótesis formadas y que quieren demostrar que son correctas.

Debería ser evidente que *cualquier error de medida sistemático sesgará las inferencias descriptivas*¹. Pensemos, por ejemplo, en el caso más simple que podamos imaginar, en el que sin darnos cuenta añadimos 1.000 dólares a la estimación de renta anual de cada una de las personas que responde a una encuesta. Es obvio que en nuestro cálculo de la renta media anual de toda la muestra habrá también una sobreestimación por ese mismo valor. Si nos interesara hacer un cálculo aproximado del efecto causal que tienen los estudios universitarios en la media de renta anual, el error de medida sistemático no tendría consecuencias sobre nuestra inferencia causal. Si, por ejemplo, el grupo de personas con estudios universitarios ganara realmente una media de 30.000 dólares, pero el grupo de control que no fue a la universidad ingresara una media de 25.000, nuestra estimación del efecto causal que tienen los estudios universitarios sobre la renta anual mostraría un incremento de 5.000 dólares. Si la renta de cada persona en ambos grupos se hubiera sobreestimado en la misma cantidad (digamos, 1.000 dólares de nuevo), nuestro efecto causal —que ahora procede de restarle 26.000 dólares a 31.000— seguiría siendo 5.000. Por lo tanto, *el error de medida sistemático que afecta a todas las unidades por medio de una misma cantidad constante no sesga la inferencia causal* (esto puede apreciarse mejor si nos centramos en la versión relativa a los efectos constantes del supuesto de homogeneidad de las unidades que se expuso en el apartado 3.1 del capítulo 3).

Sin embargo, supongamos que hubiera un error sistemático en una parte de la muestra: los licenciados universitarios exageran sistemáticamente sus ingresos porque quieren impresionar al entrevistador, pero las cantidades que señala el grupo de control son más correctas. En este caso, tanto la inferencia descriptiva como la del efecto causal que tienen los estudios sobre la renta estarían sesgadas. Si sabemos que existe este problema, podremos plantear mejores preguntas en la encuesta u obtener la información de otro modo. Si los datos ya se han recogido y no podemos recabar más, podríamos, por lo menos, determinar en qué dirección va el sesgo para corregirlo posteriormente.

Para hacer más hincapié en este asunto, pensemos en un ejemplo tomado de los escritos sobre integración regional, dentro de la bibliografía relativa a las relaciones internacionales. En esta área, más que en cualquier otra de las relaciones internacionales, se ha intentado contrastar hipótesis concretas, a veces mediante indicadores cuantitativos. Sin embargo, uno de sus conceptos más importantes —en qué medida el control de las políticas se transmite desde los estados-nación a las organizaciones internacionales— no puede expresarse fácilmente con medidas cuantitativas válidas. Por lo tanto, los investigadores han ideado formas cualitativas de medirlo y las han codificado utilizando su conocimiento profundo del problema (por ejemplo, Lindberg y Scheingold, 1970: 71, p. 71, cuadro 3.1). Han incorporado a sus variables explicativas conceptos subjetivos como la «complementariedad del valor de las élites» o la «forma de tomar decisiones» (véanse Nye, 1971, o Lindberg y Scheingold, 1971), con la intención de analizar las relaciones que se establecen entre las variables explicativas y la dependiente cuando en la medición de ambas se utilizan tales conceptos.

Este enfoque respondía a una preocupación por la validez: había investigadores expertos que codificaban la información y que podían examinar si era relevante para los conceptos subyacentes en sus medidas. Sin embargo, con este enfoque se corría el riesgo de cometer un error de medida subjetivo. Los investigadores tenían que controlarse mucho para no codificar las variables explicativas en función de sus propias posturas o expectativas teóricas. En cualquiera de los casos, puede que lo hayan hecho, pero para el lector resulta difícil determinar hasta qué punto lograron su objetivos.

En tales circunstancias, nuestro consejo es que, en primer lugar, se intenten utilizar valoraciones que hayan hecho *otros investigadores* para propósitos completamente diferentes. En la investigación cuantitativa o cualitativa, este elemento de arbitrariedad nos garantiza que las medidas no se verán influenciadas por nuestras hipótesis, las cuales, cabe suponer, se han conformado posteriormente. Esta estrategia suele seguirse en los estudios cuantitativos —el investigador toma medidas ajenas y las utiliza para sus propósitos—, pero también es muy buena para los cualitativos. Por ejemplo, se puede hacer que observadores expertos codifiquen conjuntamente ciertas variables clave para un problema mediante sus diversas interpreta-

ciones y explicaciones. Aprovechando el análisis y los conocimientos que comparten dichos informadores, se podrían organizar bancos de datos con categorías normalizadas y después utilizarlos para evaluar hipótesis. Si usted es la primera persona que usa un conjunto de variables, sería útil dejar que *otros expertos* codificaran las variables, sin conocer la teoría que usted tiene sobre la relación que quiere evaluar. Enséñeles sus notas de trabajo y las entrevistas grabadas para ver si sus conclusiones sobre las medidas son las mismas que las de usted. Como la fiabilidad de una codificación aumenta si puede reproducirse, el resultado de esta revisión de las medidas será mejor cuanto más cualificados estén los observadores que la realicen.

1.2 Error de medida no sistemático

Otro de los problemas a los que se enfrentan todos los investigadores, ya sea en los estudios cuantitativos o cualitativos, es el error de medida no sistemático². El error no sistemático no sesga la medida de la variable. En este contexto, definimos las variables que presentan errores de medida no sistemáticos, o aleatorios, como aquellas en las que se registran valores que a veces son demasiado altos y otras demasiado bajos. Es evidente que el error aleatorio produce ineficiencias, pero no sesgo cuando se extraen inferencias descriptivas. Este asunto ya se ha analizado en el apartado 7.1 del capítulo 2. Ahora vamos más allá de las consecuencias que tiene el error de medida aleatorio en la inferencia descriptiva para ocuparnos de las que produce en la causal.

Cuando se hace una estimación de efectos causales, el efecto del error de medida aleatorio es diferente si aparece en la variable explicativa o en la dependiente. Cuando lo encontramos en esta última, reduce la eficiencia de la estimación causal pero no la sesga, y puede hacer que los valores de las relaciones causales sean a veces demasiado altos y otras demasiado bajos. Sin embargo, la media de la estimación será correcta. De hecho, en la variable dependiente este tipo de error de medida no es diferente, e incluso no se distingue, del error aleatorio común que, tal como se refleja en esa variable, existe en el mundo.

En una variable explicativa, el error aleatorio también puede generar ineficiencias que conduzcan a estimaciones vacilantes, altas o bajas. Sin embargo, sus consecuencias también son diferentes a las que produce en la variable dependiente, ya que en la explicativa el error aleatorio sesga la estimación de la relación que existe entre ambas. El sesgo adopta una forma particular: hace que la estimación indique la existencia de una relación causal más débil de lo que es en realidad. Si la auténtica relación es positiva, el error aleatorio en la variable explicativa sesgará la estimación haciéndola más pequeña o nula. Si es negativa, la sesgará aumentándola hasta hacerla también nula.

Como esta diferencia entre el efecto que tiene el error aleatorio en la variable explicativa y en la dependiente no resulta evidente de forma intuitiva, ofreceremos pruebas formales de él, mostrándolo gráficamente y dando un ejemplo ilustrativo. Comenzamos por el efecto del error aleatorio en la variable dependiente.

1.2.1 Error de medida no sistemático en la variable dependiente

El error de medida *no sistemático* o aleatorio en una variable dependiente no sesga la estimación habitual del efecto causal, pero sí la hace menos eficiente. En cualquier aplicación, esta falta de eficiencia tendrá resultados impredecibles, indicando unas veces la existencia de inferencias causales demasiado grandes y otras demasiado pequeñas. Por consiguiente, este tipo de error aumenta la incertidumbre de nuestras inferencias. Dicho de otro modo, el error de medida aleatorio en una variable dependiente crea un problema similar al que produce tener un número de observaciones reducido; en ambos casos, la cantidad de información que podemos utilizar para resolver un problema es menor de lo que quisiéramos. El resultado es que *el error de medida aleatorio en la variable dependiente produce estimaciones de los efectos causales menos eficientes y más inciertas.*

Cuando usemos diversos conjuntos de datos —como habría que hacer siempre que se pueda—, las estimaciones basadas en variables dependientes con error de medida aleatorio serán inestables. Algunos conjuntos de datos nos darán pruebas de la existencia de intensas relaciones, mientras que otros presentarán efectos negativos o dirán que no hay ninguno, incluso si la verdadera relación no ha cambiado en absoluto. Esta ineficiencia dificulta, a veces considerablemente, el hallazgo de rasgos descriptivos sistemáticos o causales en un conjunto de datos o (quizás de forma más evidente) en la comparación de varios de ellos. Con frecuencia, las estimaciones de incertidumbre serán mayores que las que indica el tamaño de las relaciones que hay entre nuestras variables. Por consiguiente, puede que no tengamos suficiente información para determinar la existencia de un efecto causal que está oculto por el error aleatorio en la variable dependiente (y representado en la mayor incertidumbre de una inferencia). Los investigadores cualitativos y cuantitativos que sean conscientes de este resultado general no podrán enfrentarse a este error más que poniendo un mayor interés en mejorar las medidas de las observaciones de que disponen o recogiendo otras nuevas que tengan los mismos niveles de error de medida (o menores). Comprender lo que suponen estos resultados en un determinado número de datos hará que los investigadores puedan matizar más adecuadamente sus conclusiones, y el hecho de que se reconozca de manera tan explícita la incertidumbre quizá les induzca a ellos u a otros a llevar a cabo estudios complementarios que midan con más cuidado las variables dependientes

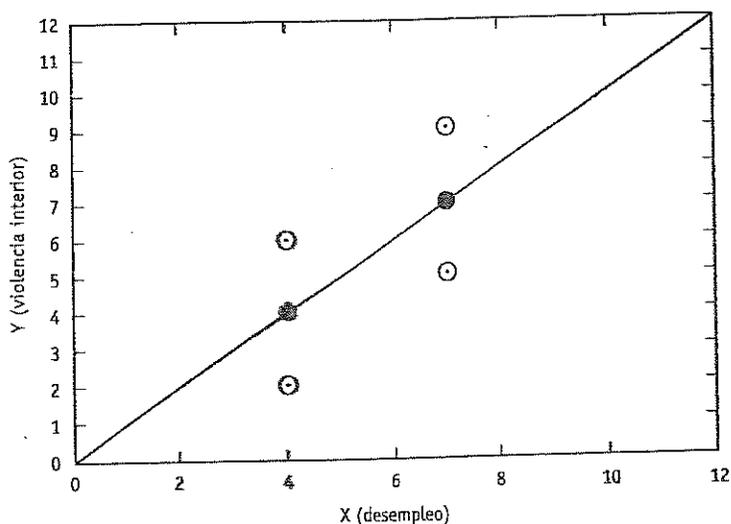
(o que tengan un mayor número de observaciones). El reconocerlo así tendría que ser todavía más útil para el diseño de la investigación, ya que los investigadores suelen tener que alcanzar un equilibrio entre la necesidad de medir con más precisión y la de obtener más observaciones. Este último objetivo es más importante para nuestra hipótesis: necesitamos determinar si la información que ésta baraja puede obtenerse mejor utilizando más observaciones de los casos actuales o si hay que recoger más datos.

Pensemos en el siguiente ejemplo de error de medida aleatorio en una variable dependiente. Al estudiar el efecto de la situación económica sobre los delitos violentos en los países en desarrollo o en las regiones de un solo país de este tipo, podemos medir la variable dependiente (violencia ilegal) observando a cada comunidad durante un corto período de tiempo. Por supuesto, estas observaciones nos darán medidas relativamente insuficientes, ya que su media será correcta, pero, en algunas comunidades, no podremos conocer muchos delitos y subestimaremos el valor medio de la violencia, mientras que en otras veremos que hay mucha criminalidad y sobreestimaremos ese mismo indicador.

Supongamos que nuestra medida de la variable explicativa —la situación económica— sea el porcentaje de desempleados de la comunidad y que la calculemos bastante bien (quizá utilizando buenos datos del gobierno). Si estudiáramos la influencia que tiene la economía —tal como la representa el porcentaje de parados— en la cantidad media de delitos violentos, cabría esperar resultados muy inciertos —también inestables si se comparan diversas aplicaciones—, precisamente porque la variable dependiente se ha medido de forma imperfecta, aunque, como promedio, la técnica de medición es correcta. El saber que ésta es la causa del problema y estar seguros de que la relación entre las variables debía ser intensa justifica suficientemente la necesidad de un nuevo estudio en el que podamos observar la delincuencia que se registra en más comunidades o durante períodos más largos. Una vez más, comprobamos que el error de medida y las pocas observaciones generan problemas similares. Podríamos mejorar la eficiencia aumentando la exactitud de nuestras observaciones (quizá recurriendo a buenas estadísticas policiales para reducir el error de medida) o incrementando el número de observaciones mal medidas en diversas comunidades. En cualquiera de estos casos, la solución es aumentar la cantidad de información que se trae a colación para solucionar el problema inferencial. Éste es otro ejemplo de por qué la cantidad de *información* que barajamos en relación con un problema es más importante que el número total de observaciones de que disponemos (siendo esta cifra la medida de la información).

Para demostrar por qué es así, utilizamos una versión simplificada de este ejemplo, primero de forma gráfica y después por medio de una prueba más formal. En el gráfico 5.1, el eje horizontal representa el desempleo. Imaginamos que las dos categorías («4%» y «7%») se han medido bien. El eje vertical representa una medida de los delitos violentos.

Gráfico 5.1 Error de medida en la variable dependiente



En el gráfico 5.1, los dos círculos negros pueden representar un estudio sencillo en el que no hay error de medida en ninguna variable. Podemos imaginar que disponemos de un gran número de observaciones y que todas coinciden con esos densos círculos negros, de manera que se conoce bastante bien la posición de ambos. Por el contrario, podemos imaginar que sólo tenemos dos observaciones y que en ellas hay poco error no sistemático de cualquier tipo. Evidentemente, no es probable que en la realidad ocurra ninguno de estos casos, pero el modelo subraya los problemas esenciales que plantea el error de medida cuando aparece en la variable dependiente de un caso más general y complicado.

Imaginemos ahora otro estudio en el que los delitos violentos se midan con un error no sistemático. Para recalcar el hecho de que la media de esas medidas es correcta, marcamos con un punto cada uno de los círculos blancos que flanquean simétricamente, por arriba y por abajo, los primeros círculos negros³. Una nueva línea que represente los seis círculos de datos será exactamente igual a la primera que hemos trazado. De nuevo, hay que señalar que esta línea se traza minimizando los errores de predicción, las desviaciones verticales de la línea.

Sin embargo, la nueva línea será más incierta en varios sentidos. Por ejemplo, los puntos estarán casi tan bien representados tanto si la línea está un poco más empinada como si es ligeramente más plana. Además, su posición vertical también será más incierta, y ella misma pronosticará peor dónde habría de estar cada uno de los círculos de datos. El resultado es que

el error de medida en la variable dependiente reduce la eficiencia de las estimaciones. Aunque éstas no tengan sesgo —es decir, en la media de numerosos estudios similares—, sí podrían estar muy alejadas de la realidad en cada uno de ellos.

1.2.2 Error de medida no sistemático en una variable explicativa

Como hemos señalado antes, el error no sistemático en la variable explicativa tiene las mismas consecuencias para las estimaciones del valor de esa variable —en las inferencias descriptivas— que para las que afectan al valor de la variable dependiente: a veces las medidas serán demasiado altas, a veces demasiado bajas, pero su media será correcta. Como ocurre con el error no sistemático en la variable dependiente, el error aleatorio en la explicativa también puede hacer que las estimaciones de los efectos causales sean inciertas e ineficientes. Sin embargo, cuando este error se produce en las variables explicativas, tiene una consecuencia bastante diferente de las que se registran cuando aparece en la dependiente. Cuando es la variable explicativa la que se mide con un error aleatorio, en las estimaciones de la relación causal aparece un sesgo sistemático que las inclina hacia un valor nulo, hacia la ausencia de relación. Dicho de otro modo, cuando hay un verdadero vínculo causal entre la variable explicativa y la dependiente, un error aleatorio en la primera puede ocultar este hecho al reducir el valor de la relación. Si queremos contrastar nuestra hipótesis en diversos conjuntos de datos, no sólo encontraremos una gran variedad de resultados, como ocurre cuando hay un error aleatorio en la variable dependiente, sino que también hallaremos un sesgo sistemático en ellos que nos inducirá a pensar que la relación que existe es más débil que la que hay en realidad.

Al igual que ocurre cuando el error de medida se produce en la variable dependiente, aunque nos demos cuenta de que tal error existe en la explicativa, sus consecuencias no se reducirán con un análisis más cuidadoso de las variables que se han medido con errores, a menos que se sigan nuestros consejos. Evidentemente, con mejores medidas mejorará la situación.

Pensemos de nuevo en nuestro estudio de los efectos del desempleo en la delincuencia que se registra en varias comunidades de un país subdesarrollado. Sin embargo, supongamos que la situación que reflejan los datos sea la opuesta a la que mencionábamos anteriormente: en el país que estamos estudiando las cifras de delincuencia son precisas y es fácil que las autoridades las proporcionen, pero el desempleo es un problema político y, en consecuencia, no puede medirse con exactitud. Como no se permiten las encuestas con muestras sistemáticas, decidimos calibrar el paro mediante la observación directa (al igual que antes hicimos con la delincuencia). Inferimos el índice de desempleo del número de personas que están

Un análisis formal del error de medida en y

Consideremos un sencillo ejemplo lineal con una variable dependiente medida con errores y una explicativa que carece de ellos. Nos interesa hacer una estimación del parámetro de efecto β :

$$E(Y^*) = \beta X$$

También especificamos un segundo rasgo de las variables aleatorias, la varianza:

$$V(Y_i^*) = \sigma^2$$

la cual presuponemos que es igual en todas las unidades $i = 1, \dots, n$.

Aunque estas ecuaciones definen nuestro modelo, por desgracia no observamos Y^* sino Y , donde

$$Y = Y^* + U$$

Es decir, que la variable dependiente observada Y es igual a la verdadera variable dependiente Y^* , más algún error de medida aleatorio, U . Para formalizar la idea de que U contiene sólo un error de medida *no sistemático*, necesitamos que éste desaparezca en la media de reproducciones hipotéticas, $E(U) = 0$, y que no se correlacione con la verdadera variable dependiente, $C(U, Y^*) = 0$, ni con la explicativa, $C(U, X) = 0$ ⁵. Además, presuponemos que el error de medida tiene una varianza $V(U_i) = \tau^2$ en cada una de las unidades i . Si τ^2 vale cero, Y no contendrá ningún error de medida y será igual a Y^* ; cuanto más grande sea la varianza, más error contendrá nuestra medida Y .

¿Cómo afecta el error de medida aleatorio en la variable dependiente a la estimación de β ? Para saberlo, utilizamos nuestro estimador habitual con Y en vez de con Y^* :

$$b = \frac{\sum_{i=1}^n Y_i X_i}{\sum_{i=1}^n X_i^2}$$

y después calculamos la media en las reproducciones hipotéticas:

$$E(b) = E \left(\frac{\sum_{i=1}^n X_i Y_i}{\sum_{i=1}^n X_i^2} \right)$$

$$\begin{aligned} &= \frac{\sum_{i=1}^n X_i E(Y_i)}{\sum_{i=1}^n X_i^2} \\ &= \frac{\sum_{i=1}^n X_i E(Y_i + U)}{\sum_{i=1}^n X_i^2} \\ &= \frac{\sum_{i=1}^n X_i^2 \beta}{\sum_{i=1}^n X_i^2} \\ &= \beta \end{aligned}$$

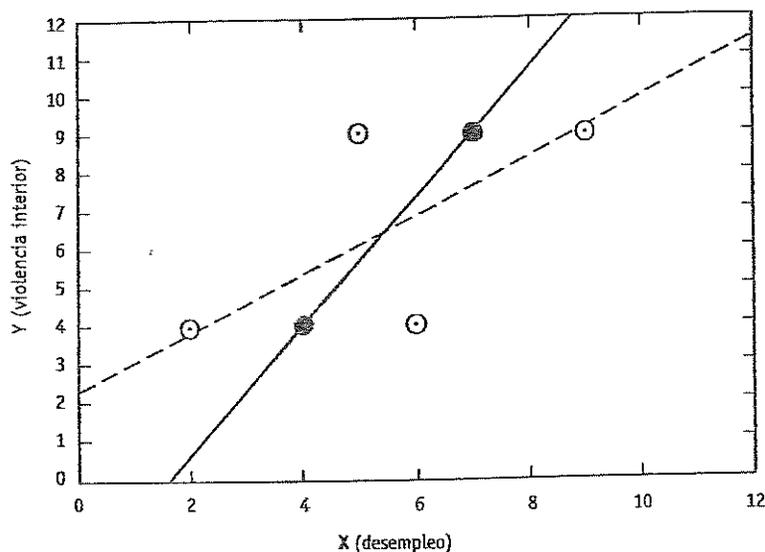
Este análisis demuestra que, incluso cuando haya un error de medida en la variable dependiente, el estimador típico no estará sesgado (será igual a la media de β), tal como mostramos que ocurría —en la ecuación (3.8)— cuando una variable dependiente no tenía error de medida.

Sin embargo, para completar este análisis, tenemos que comprobar la eficiencia de nuestro estimador en presencia de una variable dependiente que se ha medido con errores. Utilizamos el procedimiento habitual:

$$\begin{aligned} V(b) &= V \left(\frac{\sum_{i=1}^n X_i Y_i}{\sum_{i=1}^n X_i^2} \right) \quad (5.1) \\ &= \frac{1}{\left(\sum_{i=1}^n X_i^2 \right)^2} \sum_{i=1}^n X_i^2 V(Y_i^* + U) \\ &= \frac{\sigma^2 + \tau^2}{\sum_{i=1}^n X_i^2} \end{aligned}$$

Hay que señalar que aquí este estimador es *menos* eficiente que cuando se aplica a datos que carecen de error de medida en la variable dependiente (compárese con la ecuación [3.9]), por lo que respecta al nivel de error de medida que aparece en dicha variable τ^2 .

Gráfico 5.2 Error de medida en la variable explicativa



de pie sin hacer nada en el centro de varios pueblos cuando pasamos en coche. Como la hora y el día en que observamos los pueblos varían, al igual que lo hace la situación meteorológica, habrá un alto grado de error aleatorio en nuestras estimaciones del índice de paro. En un gran número de pueblos nuestros cálculos no serán sistemáticamente altos o bajos. Una estimación que se basara en cualquier pareja de pueblos sería bastante ineficiente: quizá las observaciones para cualquiera de ellas se hayan hecho en domingo (cuando puede que mucha gente esté largo rato en la calle) o en un día de lluvia (cuando pocos lo harían). Sin embargo, muchas observaciones de parejas de pueblos que se hayan hecho en diversos momentos de diferentes días, con sol o con lluvia, producirán una media correcta de las estimaciones del efecto. No obstante, como se ha señalado anteriormente, la consecuencia será muy diferente de la que producía un error similar en nuestra medida de la variable dependiente denominada delitos violentos.

El gráfico 5.2 ejemplifica esta situación. Los dos círculos negros representan un estudio en el que no hay error de medida en ninguna variable⁶. Por consiguiente, la inclinación de la línea continua representa la estimación correcta del efecto causal que tiene el desempleo sobre la delincuencia. Para mostrar las consecuencias de este error, añadimos dos círculos más (los blancos) a derecha e izquierda de cada uno de los negros, con el fin de representar un error de medida en la variable explicativa

va cuya media es correcta (o sea, igual a la del círculo negro). La línea discontinua se ajusta a los círculos blancos, y la diferencia entre las dos líneas es el sesgo debido al error de medida aleatorio que hay en la variable explicativa. De nuevo recalcamos que las líneas se trazan de manera que minimicen los errores que se generan al predecir la variable dependiente (que aparecen en el gráfico como desviaciones *verticales* de la línea que pasa por los círculos), según cada uno de los valores de las variables explicativas.

De este modo, el efecto estimado del desempleo, que aquí se ha calculado con un considerable error aleatorio, será mucho más pequeño que el verdadero (ya que la línea discontinua está menos inclinada). Como sabemos que existe un error de medida en la variable explicativa, podríamos inferir que el auténtico efecto que tiene el desempleo sobre la delincuencia es mayor que la correlación observada en este proyecto de investigación.

El análisis de las consecuencias del error de medida en una variable explicativa nos lleva a dos directrices de tipo práctico:

1. Si, en principio, un análisis indica que no hay ningún efecto, el efecto verdadero será muy difícil de calibrar, ya que no se sabe la dirección del sesgo; por consiguiente, el análisis será bastante impreciso y así habrá que describirlo. Puede que el auténtico efecto sea nulo, negativo o positivo, y que nada en los datos nos señale cómo es.
2. Sin embargo, si un análisis indica que la variable explicativa con error de medida aleatorio tiene un pequeño efecto positivo, tendremos que utilizar los resultados de este apartado como justificación para concluir que el auténtico efecto probablemente sea aún mayor que el que hemos encontrado. Del mismo modo, si lo que hallamos es un pequeño efecto negativo, los resultados de este apartado podrán utilizarse como prueba de que el efecto verdadero probablemente indique la existencia de una relación negativa aún mayor.

Como el error de medida es un rasgo fundamental de todos los estudios cualitativos, estas directrices tendrían que poder aplicarse a todos ellos.

Es preciso matizar de alguna forma estas conclusiones para que los investigadores sepan exactamente cuándo son aplicables y cuándo no lo son. En primer lugar, el análisis del recuadro inferior, en el que se basan nuestras recomendaciones, puede aplicarse a modelos en los que sólo exista una variable explicativa. Hay resultados similares que sí son aplicables a muchas situaciones en las que existen múltiples variables explicativas, pero no a todas. Del mismo modo, el análisis también es aplicable si el investigador tiene muchas variables explicativas, pero sólo una con un considerable error de medida. Sin embargo, si se están analizando de forma simultánea las consecuencias de muchas variables y cada una de ellas padece formas diferentes de error de medida, sólo podremos calibrar los ti-

pos de sesgo que pueden aparecer si extendemos el análisis formal que aparece más abajo. Resulta que, a pesar de que los cualitativistas suelen tener muchas variables explicativas, lo más habitual es que estudien el efecto de cada una de ellas de forma secuencial más que simultánea. Por desgracia, como se describe en el apartado 2, este procedimiento puede causar problemas como el del sesgo de la variable omitida, pero hace que resultados similares a los analizados aquí sean bastante aplicables en la investigación cualitativa.

2. Exclusión de variables relevantes: el sesgo

La mayoría de los cualitativistas se dan cuenta de la importancia de controlar los posibles efectos falsos de otras variables cuando se hace una estimación de las consecuencias que tiene una variable sobre otra. Entre las formas de ejercer este tipo de control se encuentran los métodos de la diferencia y de la similitud de John Stuart Mill (1843) (a los que, irónicamente, Przeworski y Teune [1982] hacen referencia calificándolos, respectivamente, como los sistemas de diseño más similares y más diferentes); la «comparación de casos disciplinada-configuradora» de Verba (1967) (que es parecida a la «comparación estructurada-centrada» de George [1982]), así como diversas formas de utilizar supuestos en los que se mantienen casi todos los rasgos constantes y otros contrafácticos parecidos. Estas expresiones se mencionan con frecuencia, pero a los investigadores les suele resultar difícil aplicarlas de manera eficaz. Por desgracia, los cualitativistas disponen de pocas herramientas para expresar las consecuencias precisas de no tener en cuenta más variables en determinados momentos de una investigación, o sea, de ocuparse del «sesgo de la variable omitida». En este apartado, aportamos esas herramientas.

Comenzamos por analizar este problema explicando con palabras las consecuencias del sesgo de la variable omitida y continuamos después con un análisis formal. Posteriormente, nos centraremos en cuestiones más amplias del diseño de investigación que suscita este tipo de sesgo.

2.1 Cómo calibrar el sesgo de las variables omitidas

Supongamos que queremos hacer una estimación del efecto causal que tiene nuestra variable explicativa X_1 sobre la dependiente Y . Si estamos haciendo un análisis cuantitativo, este efecto causal de X_1 sobre Y lo expresaremos mediante β_1 , que se puede estimar mediante una ecuación de regresión β_1 , o utilizando otra forma de análisis que nos dé una estimación b_1 de β_1 . Si estamos realizando investigación cualitativa también intentaremos obtener esa estimación del efecto causal; sin embargo, en este caso, el

cálculo dependerá de un argumento verbal y de la valoración del investigador, que se basan en la experiencia y el criterio.

Supongamos que después de haber hecho las estimaciones (cuantitativa o cualitativamente) un colega eche un vistazo a nuestro análisis y plantee la objeción de que hemos omitido una importante variable de control, X_2 . Hemos estado haciendo una estimación del efecto que tienen los gastos en las campañas electorales sobre la proporción de voto que recibe un candidato al Congreso estadounidense. Para este colega, puede que nuestra conclusión sea falsa por el «sesgo de la variable omitida». Es decir, lo que sugiere es que nuestra estimación b_1 de β_1 es incorrecta porque no hemos tenido en cuenta otra variable explicativa X_2 (como medir si los candidatos son ya congresistas). El verdadero modelo tendría que controlar el efecto de la nueva variable.

¿Cómo podemos evaluar la afirmación de nuestro colega? Especialmente, ¿en qué condiciones influiría sobre nuestra estimación de las consecuencias que tiene el gasto en los votos el haber omitido la variable que calibra el hecho de ostentar ya el cargo de congresista, y en qué condiciones no influiría? Es evidente que omitir un término que mida la influencia de ser ya congresista no tendrá importancia si este hecho no influye en la variable dependiente; o sea, que si X_2 es irrelevante, porque no tiene ningún efecto sobre Y , no producirá sesgo. Éste es el primer caso especial: las variables omitidas irrelevantes no producen sesgos. De este modo, si el hecho de estar ya en el poder careciera de consecuencias electorales, podríamos olvidarnos de haberlo omitido.

El segundo caso especial, que tampoco produce sesgos, se da cuando la variable omitida no se correlaciona con la explicativa incluida. De este modo, tampoco hay sesgo si el hecho de ser ya congresista no se correlaciona con nuestra variable explicativa, el gasto en la campaña electoral. De forma intuitiva, podemos decir que cuando una variable omitida no se correlaciona con la principal variable explicativa que nos interesa, someterla a control no cambiará la estimación del efecto causal que tiene sobre nuestra variable principal, porque lo que controlamos es la parte de variación que las dos variables tienen en común, si es que hay alguna. Por lo tanto, *podemos omitir variables de control con tranquilidad, aunque tengan una gran influencia en la variable dependiente, siempre que no cambien con la variable explicativa incluida en el análisis*⁷.

Si estos casos especiales no se sostienen a causa de alguna variable omitida (por ejemplo, si ésta se correlaciona con la variable explicativa incluida y tiene consecuencias en la dependiente), el no controlarla sesgará nuestra estimación (o percepción) del efecto de la variable incluida. En el caso que nos ocupa, las críticas de nuestro colega serían correctas, ya que el hecho de ser ya congresista está relacionado con la variable dependiente y también con la independiente: los que ya están en el poder reciben más votos y gastan menos.

Un análisis formal del error de medida aleatorio en X

Definimos primero un modelo como sigue:

$$E(Y) = \beta X^*$$

en el que no se observa la verdadera variable explicativa X^* sino que se ve X , donde

$$X = X^* + U$$

y el error de medida aleatorio U tiene propiedades similares a las de antes: su media es cero, $E(U) = 0$, y no se correlaciona ni con la verdadera variable explicativa, $C(U, X^*) = 0$, ni con la dependiente, $C(U, Y) = 0$.

¿Qué ocurre cuando utilizamos el estimador típico de β con una X plagada de errores, en vez de con la X^* no observada? Esta situación es habitual en la investigación cualitativa si hay un error de medida y no se hace ningún ajuste especial en los resultados que se obtienen. Para analizar el efecto de este procedimiento evaluamos el sesgo, que va a ser la principal consecuencia de este tipo de problema de medida. Por lo tanto, comenzamos por aplicar el estimador típico de la ecuación (3.7) a las X e Y observadas en el modelo anterior.

$$\begin{aligned} b &= \frac{\sum_{i=1}^n Y_i X_i}{\sum_{i=1}^n X_i^2} & (5.2) \\ &= \frac{\sum_{i=1}^n (X_i^* + U_i) Y_i}{\sum_{i=1}^n (X_i^* + U_i)^2} \\ &= \frac{\sum_{i=1}^n X_i^* Y_i + \left(\sum_{i=1}^n U_i Y_i \right)}{\sum_{i=1}^n X_i^{*2} + \sum_{i=1}^n U_i^2 + \left(2 \sum_{i=1}^n X_i^* U_i \right)} \end{aligned}$$

Esta idea puede plantearse en términos formales, centrándose en la última línea de la ecuación (5.5) que aparece en el recuadro inferior:

$$E(b_1) = \beta_1 + F\beta_2 \quad (5.3)$$

Debería quedar claro que b estará sesgada, $E(b) \neq \beta$. Además, la media de los dos términos que están entre paréntesis en la última línea de la ecuación (5.2) será cero porque presuponemos que U e Y , así como U y X^* , no se correlacionan (o sea, que $C(U, Y) = E(U, Y) = 0$). Por lo tanto, esta ecuación se reduce aproximadamente a⁸:

$$= \frac{\sum_{i=1}^n X_i^* Y_i}{\sum_{i=1}^n X_i^{*2} + \sum_{i=1}^n U_i^2}$$

Esta ecuación con el estimador de β aplicado al modelo anterior sólo se diferencia de la que utiliza el estimador típico por el término adicional que

hay en el denominador, $\sum_{i=1}^n U_i^2$ (compárese con la ecuación [3.7]), que representa el grado de error de medida en X , la varianza muestral del error U . Si no hay error de medida, este término será nulo y la ecuación se reducirá al estimador típico de la (3.7), ya que en realidad habríamos observado los auténticos valores de la variable explicativa.

En el caso general en el que hay algún error de medida $\sum_{i=1}^n U_i^2$ será una suma de términos cuadráticos y, por tanto, siempre será positiva. Como este término se suma al denominador, b se acercará a cero. Si el estimador correcto produjera un número positivo alto, el error de medida aleatorio en la variable explicativa haría que el investigador pensara erróneamente que b tiene un valor positivo pero más pequeño. Si la estimación que se basa en X^* tuviera un valor alto negativo, un investigador que analizara datos con error de medida aleatorio pensaría que dicha estimación es un número negativo más pequeño.

Sería fácil utilizar este análisis formal para demostrar que el error de medida aleatorio en las variables explicativas también causa ineficiencias, pero el sesgo suele ser un problema más grave y nos ocuparemos de él en primer lugar.

Ésta es la ecuación que se utiliza para calcular el sesgo en la estimación del efecto que tiene X_1 sobre la variable dependiente Y . En esta ecuación, F representa el grado de correlación que hay entre las dos variables explicativas, X_1 y X_2 ⁹. Si el estimador que se calcula utilizando sólo X_1 como variable explicativa (b_1) no estuviera sesgado, su media sería igual a β_1 ; o sea, que se podría decir que $E(b_1) = \beta_1$. Este estimador no está sesgado en nin-

gundo de los dos casos especiales en los que el término que expresa el sesgo, $F\beta_2$, es igual a cero. Resulta fácil apreciar que esto formaliza las condiciones de falta de sesgo que antes hemos señalado. Es decir, que podemos omitir una variable de control si:

- La variable omitida no tiene un efecto causal sobre la dependiente (o sea, $\beta_2 = 0$, independientemente del tipo de relación que exista entre la variable incluida y la excluida, F).
- La variable omitida no se correlaciona con la incluida (es decir, $F = 0$ cualquiera que sea el valor de β_2).

Cuando sospechemos que hay una variable omitida que puede estar sesgando los resultados, nuestro análisis no podrá terminar aquí. Si es posible, hay que controlar esa variable omitida y, aunque no podamos hacerlo por falta de una buena fuente de datos acerca de ella, nuestro modelo podrá ayudarnos a determinar la dirección del sesgo, lo cual puede ser muy útil. Una sobreestimación puede reforzar sustancialmente un argumento existente y una subestimación puede debilitarlo.

Supongamos, por ejemplo, que estamos estudiando unos pocos estados subsaharianos y que nos damos cuenta de que los golpes de Estado son más frecuentes en regímenes que practican la represión política (aquellos en los que β_1 , el efecto que tiene la represión en la probabilidad de golpe, es positivo). Esto significa que la variable explicativa es el grado de represión política y que la dependiente corresponde a la probabilidad de que haya un golpe de Estado. La unidad de análisis son los países africanos del área subsahariana. Incluso, podríamos incorporar a la muestra otros países africanos y llegar a la misma conclusión. Sin embargo, supongamos que no hemos considerado los efectos que puede tener la situación económica en los golpes de Estado. Aunque quizá no dispongamos de datos sobre este indicador, sería razonable plantear la hipótesis de que probablemente el desempleo incrementa la probabilidad de golpe ($\beta_2 > 0$), y también parece plausible que haya una correlación positiva entre desempleo y represión política ($F > 0$). También presuponemos, porque es útil en este ejemplo, que las condiciones económicas son previas a nuestra variable causal clave: el nivel de represión política. Si así fuera, el grado de sesgo de nuestro análisis podría ser considerable. Como el desempleo se correlaciona positivamente con la variable dependiente y con la explicativa ($F\beta_2 > 0$ en este caso), excluir ese factor supondría que, sin darnos cuenta, estaríamos haciendo una estimación del efecto que tienen la represión y el desempleo en la posibilidad de que haya un golpe de Estado, en vez de hacerla únicamente de la represión ($\beta_1 + F\beta_2$ en vez de β_1). Además, al ser el impacto conjunto de ambos factores mayor que el de la represión por sí sola, ($\beta_1 + F\beta_2$ es mayor que β_1), la estimación del efecto medio de esta última variable (b_1) sería demasiado grande. Por lo tanto, el análisis demuestra que, al excluir las consecuencias del desempleo, sobreestimamos las que tie-

ne la represión política (el impacto de esta exclusión no es el mismo que el del error de medida en las variables explicativas, ya que el sesgo de la variable omitida a veces puede tener como consecuencia que se haga una estimación positiva de una relación negativa).

Omitir variables relevantes no siempre redundará en una sobreestimación de los efectos causales. Por ejemplo, sería razonable hacer una hipótesis que señalara que en otros países (quizá esto sería tema para otro estudio) la represión política y el desempleo tienen una relación inversa (F es negativo). En estos países, la represión política podría permitir al gobierno controlar a las facciones en lucha, imponer la paz desde arriba y dar trabajo a la mayoría de la población. A su vez, esta situación conllevaría que el efecto del sesgo que introduce la relación negativa entre desempleo y represión ($F\beta_2$) también sería negativo, siempre que siguiéramos estando dispuestos a presuponer que con un mayor desempleo aumentara la probabilidad de golpe en esos países. La principal consecuencia es que el efecto estimado de la represión en la probabilidad de golpe [$E(b_1)$] ahora sería menor que el efecto real (β_1). De este modo, si se excluye la situación económica, b_1 tenderá a subestimar el efecto de la represión política. Si F es lo suficientemente negativo y β_2 lo suficientemente grande, podríamos considerar que β_1 siempre será negativo y ¡llegar a la conclusión incorrecta de que una mayor represión política disminuye la posibilidad de golpe de Estado! Aunque no dispusiéramos de información suficiente sobre las tasas de desempleo para incluirla en el estudio original, un análisis como éste seguiría ayudándonos a llegar a conclusiones razonables e importantes.

Estos ejemplos tendrían que dejar claro que, en realidad, no es preciso hacer una regresión para calcular parámetros, ni para determinar el grado y dirección del sesgo, ni para llegar a conclusiones relativas a estos asuntos. Las estimaciones de tipo cualitativo e intuitivo padecen el mismo tipo de sesgos que las estrictamente cuantitativas. Este apartado demuestra que, en ambas situaciones, la información que no se refiere a los datos existentes puede ser de gran ayuda para hacer una estimación del grado de sesgo y de su dirección.

Si pensamos que quizá nuestro diseño de investigación se ha visto perjudicado por la omisión de ciertas variables pero no sabemos cuáles son, es muy posible que tengamos conclusiones imperfectas (que probablemente, en el futuro, otros investigadores detectarían). Los incentivos para profundizar más en el problema son evidentes. Por fortuna, en la mayoría de los casos, los investigadores disponen de bastante información sobre las variables que están fuera de su análisis. A veces, la información es detallada, pero sólo se refiere a ciertas unidades secundarias; es parcial pero aplicable al conjunto de la investigación, o incluso procede de estudios anteriores. Toda información, cualquiera que sea su procedencia y aunque sea incompleta, puede ayudarnos a determinar el posible nivel de sesgo que hay en nuestros efectos causales y la dirección que sigue.

Es evidente que, incluso los investigadores que comprenden las consecuencias del sesgo de la variable omitida, pueden encontrarse con dificultades al identificar las variables que han podido eliminar de sus análisis. No hay ninguna receta para enfrentarse a este problema, pero sí recomendamos a todos los investigadores, ya sean cuantitativistas o cualitativistas, que busquen sistemáticamente variables de control omitidas y que piensen si hay que incluirlas en sus estudios. En este apartado indicamos algunas directrices para proceder a esa revisión.

Las variables omitidas pueden producir dificultades aunque se disponga de información adecuada sobre las variables relevantes. A veces los investigadores tienen esa información y, creyendo que ciertas variables tienen una relación positiva con la variable dependiente, hacen una estimación secuencial de los efectos causales de dichas variables, en análisis «bivariantes» separados. Recurrir a este enfoque resulta especialmente tentador en estudios que tienen un pequeño número de observaciones, ya que, como señalamos en el apartado 1 del capítulo 4, utilizar simultáneamente muchas variables explicativas genera estimaciones poco exactas e incluso puede llevarnos a un diseño de investigación impreciso. Sin embargo, por desgracia, cada uno de los análisis excluye las otras variables relevantes, y esto hace que aparezca el sesgo de la variable omitida en cada estimación. La solución ideal no sólo consiste en recabar información sobre las variables relevantes, sino en controlar clara y *simultáneamente* todas ellas. El cualitativista debe admitir que no tener en cuenta todas las variables relevantes a la vez produce inferencias sesgadas. Reconocer las causas del sesgo es importante, aunque sea imposible eliminarlo porque se dispone de un pequeño número de observaciones.

Sin embargo, la preocupación por el sesgo de la variable omitida *no* debería llevarnos a incluir de manera automática cualquier variable cuya omisión pudiera causar sesgo porque se correlaciona con la variable independiente y tiene consecuencias sobre la dependiente. En general, *no tenemos que controlar una variable explicativa que sea, en parte, consecuencia de nuestra variable causal clave.*

En el siguiente ejemplo, suponemos que nos interesa el efecto causal que tiene aumentar el valor de la renta (variable de tratamiento) en 10.000 dólares sobre la probabilidad de que un ciudadano vote, en los Estados Unidos, a un candidato demócrata (variable dependiente). ¿Acaso habría que controlar el hecho de que esta persona, cinco minutos antes de llegar al colegio electoral, afirme en una entrevista que va a votar a los demócratas? Esta variable de control sí que influye en la dependiente y es probable que se correlacione con la explicativa. Intuitivamente, la respuesta a la pregunta anterior es no, ya que, si realmente controláramos esta variable, el efecto estimado que tendría la renta en el hecho de votar por los demócratas se atribuiría casi por completo a la variable de control, que, en este caso, apenas puede considerarse como una explicación causal alternativa.

Si aplicáramos ciegamente las normas antes señaladas para evitar el sesgo de la variable omitida, podríamos llegar a cometer el error de controlar esta variable. Después de todo, esta posible variable de control sí tiene un efecto en la dependiente (votar a los demócratas) y se correlaciona con la explicativa clave (la renta). Sin embargo, al incluirla, estaríamos atribuyendo a la variable de control parte del efecto causal que tiene la explicativa clave.

Para dar otro ejemplo, supongamos que nos interesa el efecto causal que tiene un drástico incremento de los precios del crudo sobre la actitud de la opinión pública ante la existencia de escasez de energía. Podríamos conseguir los precios del petróleo (variable causal clave) en los periódicos y utilizar las encuestas de opinión como variable dependiente para calibrar cómo percibe la opinión pública la posibilidad de que haya escasez de energía. Sin embargo, cabría preguntarse si hay que controlar la influencia que tiene la información televisiva que se refiere a los problemas energéticos. Es cierto que la atención que presta la televisión a estos asuntos se correlaciona tanto con la variable explicativa incluida (el precio del crudo) como con la dependiente (la opinión pública en relación a la escasez de energía). No obstante, como la información televisiva es, en parte, consecuencia de los precios del crudo en el mundo real, no deberíamos controlarla al valorar la influencia causal que tienen los precios del crudo en la actitud de la opinión pública ante la escasez de energía. Si, en lugar de esto, nos interesara el efecto causal de la información televisiva, habría que controlar los precios del crudo, ya que éstos son *anteriores* a la variable explicativa clave (que ahora es la información televisiva)¹⁰.

En consecuencia, para hacer una estimación del efecto total que tiene una variable explicativa, hay que enumerar todas aquellas variables que, según nuestro modelo teórico, puedan generar la dependiente. Para repetir lo que señalamos anteriormente: en general, no tenemos que controlar una variable explicativa que sea, en parte, consecuencia de la variable causal clave. Una vez que se han eliminado estas posibles variables explicativas, hay que controlar otras del mismo tipo que pudieran causar sesgo de la variable omitida (las que se correlacionan con la variable dependiente y con las explicativas incluidas en el análisis)¹¹.

La idea de que no tenemos que controlar las variables que sean consecuencia de las variables explicativas clave tiene una importante implicación para el papel de la teoría en nuestro diseño de investigación. Si pensamos en este problema, entenderemos por qué hay que partir de un modelo con fundamento teórico —o al menos tender hacia él— en vez de «escarbar en busca de datos», es decir, realizar regresiones o análisis cualitativos con cualquier variable explicativa que se nos ocurra. Sin un modelo teórico no podremos decidir qué posibles variables explicativas hay que incluir en nuestro análisis. De hecho, si no hay modelo, se pueden lograr los resultados más sólidos utilizando una variable explicativa trivial —como la inten-

ción de votar a los demócratas que se expresa cinco minutos antes de entrar al colegio electoral— y controlando todos los demás factores que se correlacionan con ella. Si no disponemos de un modelo con fundamento teórico, no podremos determinar si hay que controlar las posibles variables explicativas o hacer caso omiso de las que se correlacionan entre sí y, además, nos arriesgaremos a incurrir en el sesgo de la variable omitida o a elaborar un diseño de investigación trivial.

No es en absoluto sencillo decidir cuándo hay que incorporar más variables explicativas a nuestro análisis. El número de variables adicionales es siempre ilimitado, pero nuestros recursos no, y, sobre todo, cuantas más variables de este tipo incluyamos, menos control tendremos sobre las estimaciones de cada uno de los efectos causales. Evitar el sesgo de la variable omitida justifica la incorporación de más variables explicativas. Si se omiten variables relevantes, nuestra capacidad de hacer estimaciones correctas de las inferencias causales se verá limitada.

2.2 Ejemplos de sesgo de la variable omitida

En este apartado, damos algunos ejemplos cuantitativos y cualitativos, unos hipotéticos y otros tomados de investigaciones reales. Por ejemplo, el nivel de estudios es uno de los mejores predictores de la participación política. Cuanto mayor es, más alta es la probabilidad de que se vote y de que se participe de otras maneras en política. Supongamos que así parece ser en un nuevo conjunto de datos, pero que queremos profundizar en este asunto y ver si la relación entre las dos variables es causal y, si lo es, determinar de qué manera los estudios conducen a la participación.

Lo primero que podríamos hacer es comprobar si se han omitido variables previas a los estudios que se correlacionen con este factor y que, al mismo tiempo, produzcan participación. Dos ejemplos de ello podrían ser el compromiso político de los padres de cada individuo y la raza de éste. Puede que los padres activos en política hayan inculcado a sus hijos el interés por participar en ella y, al mismo tiempo, puede que sean la clase de padres que favorecen el éxito en los estudios. Si no incluyéramos esta variable, podríamos obtener un relación falsa entre estudios y actividad política o hacer una estimación demasiado grande de ella.

La raza podría tener el mismo papel. En una sociedad en la que hay discriminación racial, los negros pueden verse excluidos tanto de las oportunidades educativas como de la participación política. En tal caso, no sería real el aparente efecto que tienen los estudios sobre la participación. Lo mejor sería que intentáramos eliminar todas las posibles variables omitidas que justifiquen en parte o totalmente la relación que existe entre los estudios y la participación.

Un análisis formal del sesgo de la variable omitida

Comencemos con un modelo sencillo en el que hay dos variables explicativas:

$$E(Y) = X_1\beta_1 + X_2\beta_2 \quad (5.4)$$

Supongamos ahora que nos encontramos un importante análisis en el que se informa del efecto que tiene X_1 sobre Y sin controlar X_2 . ¿En qué circunstancias habría razones para criticar este estudio o justificación para buscar fondos con los que repetirlo? Para contestar a esta pregunta, evaluamos formalmente el estimador con la variable de control omitida.

El estimador de β_1 en el que eliminamos X_2 es:

$$b_1 = \frac{\sum_{i=1}^n X_{1i} Y_i}{\sum_{i=1}^n X_{1i}^2}$$

Para evaluar este estimador, tomamos el valor esperado de b_1 en varias hipotéticas reproducciones, siguiendo el modelo de la ecuación (5.4):

$$\begin{aligned} E(b_1) &= E\left(\frac{\sum_{i=1}^n X_{1i} Y_i}{\sum_{i=1}^n X_{1i}^2}\right) & (5.5) \\ &= \frac{\sum_{i=1}^n X_{1i} E(Y)}{\sum_{i=1}^n X_{1i}^2} \\ &= \frac{\sum_{i=1}^n X_{1i} (X_{1i}\beta_1 + X_{2i}\beta_2)}{\sum_{i=1}^n X_{1i}^2} \\ &= \frac{\sum_{i=1}^n X_{1i}^2 \beta_1 + \sum_{i=1}^n X_{1i} X_{2i} \beta_2}{\sum_{i=1}^n X_{1i}^2} \\ &= \beta_1 + F\beta_2 \end{aligned}$$

donde $F = \frac{\sum_{i=1}^n X_{1i} X_{2i}}{\sum_{i=1}^n X_{1i}^2}$, el coeficiente de la pendiente en la regresión de X_1 y

X_2 . La última línea de esta ecuación se reproduce en el texto de la (5.3) y ya se ha analizado con cierto detalle anteriormente.

Sin embargo, el hecho de que la relación entre ambos elementos disminuya o desaparezca cuando controlamos una variable previa no tiene por qué suponer que los estudios sean irrelevantes. Supongamos que nos damos cuenta de que el vínculo entre los estudios y la participación se debilita cuando controlamos la raza. Una de las razones podría ser que, como en el ejemplo anterior, la discriminación de los negros implique que la raza está relacionada, por un lado, con el éxito académico y, por otro, con la participación. En estas condiciones, no habría ningún vínculo causal real entre los estudios y la participación. Por otra parte, la raza podría influir en la participación política a través de los estudios, ya que la discriminación racial podría reducir el acceso de los negros a la educación. A su vez, ésta podría ser el principal factor a la hora de generar participación. En este caso, la merma de la relación entre educación y participación que se produce cuando el investigador incluye la raza en el análisis no quita importancia a la educación. En realidad, explica cómo raza y educación interactúan para influir en la participación.

Hay que señalar que estas dos situaciones son fundamentalmente diferentes. Si la menor participación de los negros fuera debida a la falta de estudios, cabría esperar que aumentara al registrarse un incremento del nivel educativo medio. Sin embargo, si la razón de la menor participación fuera la discriminación política directa que impide que los negros ejerzan sus derechos como ciudadanos, la mejora de la educación sería irrelevante para la transformación de las pautas de participación.

También podríamos buscar variables que actúen a la vez que la educación o posteriormente. Se podrían buscar variables omitidas que muestren que la relación entre estudios y participación es falsa u otras que nos ayuden a explicar de qué manera la primera favorece a la segunda. En la primera categoría podría haber una variable como la del nivel de inteligencia general del individuo (la cual podría hacer que sacara buenas notas en el colegio y llevarle a la actividad política). En la segunda categoría podría haber variables que midieran aspectos educativos como la asistencia a cursos de educación cívica, las oportunidades de formar parte de órganos de gobierno estudiantiles o el aprendizaje de técnicas básicas de comunicación. Si se apreciara que la introducción en el análisis de una o más variables del segundo grupo disminuye la relación entre éxito en los estudios y participación (cuando se controlaron las técnicas de comunicación no se apreció que el éxito en los estudios tuviera un efecto independiente sobre la participación), esto no significaría que los estudios fueran irrelevantes. Las técnicas de comunicación necesarias se aprenden en la escuela y son diferentes en cada nivel educativo. Lo que el análisis nos diría es cómo influye la educación en la participación.

Una vez más, todos estos ejemplos ponen de manifiesto por qué es necesario evaluar con un modelo teórico en mente. No hay otra manera de elegir qué variables vamos a utilizar en nuestro análisis. Una teoría que se ocupara

de cómo influyen los estudios en las actividades cívicas nos indicaría qué variables hay que incluir. Aunque en una investigación cualitativa no incorporaremos más variables a una ecuación de regresión, la lógica es bastante parecida cuando decidimos qué otros factores vamos a tener en cuenta. Pensemos en una pregunta que planteamos anteriormente: qué impacto tienen los encuentros en la cumbre sobre la cooperación entre las superpotencias. Supongamos que nos diéramos cuenta de que, entre los Estados Unidos y la URSS, este tipo de cooperación hubiera sido mayor en los años posteriores a una cumbre que en los precedentes. ¿Cómo sabríamos que ese efecto es real y que no procede de alguna variable omitida? Y, si estamos seguros de que es real, ¿cómo podríamos profundizar en su funcionamiento?

Podríamos considerar variables precedentes, relacionadas con la posibilidad de que se celebrara una cumbre y que también pudieran ser causantes directas de la cooperación. Quizá cuando los mandatarios de cada país se fían el uno del otro, las reuniones son frecuentes y los países cooperan, o puede que cuando las ambiciones geopolíticas de ambos bandos se ven limitadas por razones políticas internas, programen reuniones y cooperen. En estas circunstancias, las cumbres, en sí mismas, no favorecerían directamente la cooperación, aunque programar una podría ser un buen indicador de que las cosas van bien entre las superpotencias. También es posible que las reuniones en la cumbre sean parte de una secuencia causal, al igual que la raza podía haber influido en el nivel educativo y éste, a su vez, en la participación. Cuando los mandatarios de las superpotencias confían el uno en el otro, convocan una cumbre para reforzar esa confianza mutua, y esa reunión, a su vez, conduce a la cooperación. En este caso, la cumbre no es en absoluto irrelevante, ya que sin ella habría menos cooperación. Supongamos que tenemos en cuenta estos factores y que encontramos que las cumbres parecen tener un papel independiente: es decir, que cuando controlamos la confianza mutua previa que hay entre los mandatarios y sus ambiciones geopolíticas, la conclusión es que una cumbre parece conducir a una mayor cooperación. Aún podríamos profundizar más y preguntarnos cómo ocurre esto. Sería posible comparar las cumbres a partir de las características que pueden hacerlas más o menos exitosas y ver si esos factores tienen alguna relación con el grado de cooperación posterior. De nuevo, es preciso seleccionar los elementos que hay que tener en cuenta, y entre ellos se pueden incluir: el grado de preparación, si los problemas tratados fueron económicos más que de seguridad, el nivel de armonía interna dentro de cada nación, la situación meteorológica durante la cumbre y también la comida. La teoría habrá de guiarnos; dicho de otro modo, necesitaremos tener una idea sobre conceptos y relaciones que nos lleve a variables explicativas relevantes y que plantee hipótesis lógicas que estén en consonancia con la experiencia que se tenga de sus efectos.

A los investigadores que manejen un pequeño número de observaciones les resultará muy difícil evitar el sesgo de la variable omitida. En esta situa-

ción, la ineficiencia tiene muchos costes, entre los que se puede incluir el riesgo de que la presencia de demasiadas variables de control irrelevantes haga que el diseño de investigación sea impreciso (apartado 1 del capítulo 4). Sin embargo, omitir variables de control relevantes puede producir sesgo, y, en principio, quizá el investigador no sepa si una posible variable es relevante o no.

En este punto, podríamos caer en la tentación de afirmar que la inferencia causal es imposible con un número escaso de observaciones. Sin embargo, para nosotros, las lecciones que hay que sacar de todo esto son más limitadas y optimistas. Comprender la dificultad que comporta extraer inferencias causales válidas a partir de pocas observaciones tendría que hacernos más cautelosos a la hora de hacer aseveraciones causales. Como indicamos en el capítulo 2, las buenas descripciones e inferencias descriptivas tienen más valor que las inferencias causales erróneas. De hecho, gran parte de la investigación cualitativa mejoraría si se prestara más atención a las inferencias descriptivas válidas y menos ímpetu a la extracción de afirmaciones causales a partir de datos inadecuados y sin calibrar adecuadamente la incertidumbre. No obstante, es posible avanzar un poco en la comprensión de los problemas causales si los problemas teóricos que nos preocupan se plantean con la suficiente claridad y se vinculan a las consecuencias observables adecuadas. Una investigación reciente del área de las relaciones internacionales puede ayudarnos a incidir en este punto.

Lo que llevó a Helen Milner a escribir *Resisting Protectionism* (1988) fue una perplejidad: ¿por qué la política comercial de los Estados Unidos era más proteccionista en los años veinte que en los setenta, a pesar de los muchos parecidos que había entre ambos períodos? Su hipótesis era que la interdependencia internacional había aumentado entre una y otra década y que esto ayudaba a explicar la diferencia de comportamiento de los Estados Unidos. Sin embargo, en este nivel de análisis agregado, sólo disponía de las dos observaciones que le habían sorprendido, y éstas no podían ayudarle a distinguir entre su hipótesis y otras muchas explicaciones posibles que se podían dar a la variación observada. Por lo tanto, el grado de incertidumbre de su teoría habría sido mucho mayor si se hubiera detenido aquí. En consecuencia, tuvo que buscar en otro sitio más consecuencias observables de su teoría.

Lo que Milner hizo fue describir el proceso que pensaba que producía su efecto causal. Planteó la hipótesis de que la interdependencia económica entre las democracias capitalistas afectaba a las preferencias nacionales porque influía en las opciones de los sectores industriales y empresas, los cuales, a base de presionar, conseguían que se llevaran a cabo las políticas que querían. En consecuencia, Milner estudió diversos sectores industriales estadounidenses de los años veinte y setenta, así como otros franceses en este segundo período, y se dio cuenta de que aquellos que más invertían a escala multinacional y dependían más de las exportaciones eran los menos protec-

cionistas. Estas conclusiones le ayudaron a confirmar su teoría general acerca de las diferencias que había, en conjunto, en la política estadounidense de los dos períodos. Por lo tanto, su procedimiento fue coherente con la parte principal de nuestros consejos metodológicos: hay que especificar cuáles son las consecuencias observables de la teoría, aunque no se refieran a nuestra preocupación principal, y diseñar la investigación de manera que se puedan extraer inferencias causales sobre tales consecuencias y que sea posible utilizarlas para evaluar la teoría. De ahí que el estudio de Milner sea ejemplar en muchos sentidos.

En cuanto al diseño de investigación, el problema más importante al que se enfrentó Milner era el relativo a las posibles variables omitidas. La variable de control más evidente era el grado de competencia de las importaciones extranjeras, ya que si éste es grande suele producir tendencias más proteccionistas en las empresas. Es decir, era probable que la competencia de las importaciones se correlacionara con la variable dependiente de Milner y que, en la mayoría de los casos, fuera anterior a sus variables explicativas o que ocurriera a la vez que ellas. Si esta variable de control también se correlacionaba con sus variables explicativas clave —la inversión a escala multinacional y la dependencia respecto a las exportaciones—, sus resultados habrían estado sesgados. De hecho, según los principios de la ventaja competitiva, habría sido probable una correlación negativa entre la competencia de las importaciones y la dependencia respecto a las exportaciones, de manera que este sesgo hipotético se hubiera materializado si el primero de estos dos factores no se hubiera utilizado como control.

Milner se enfrentó a este problema seleccionando para su estudio únicamente aquellos sectores industriales a los que afectaba de manera determinante la competencia extranjera. En consecuencia, mantuvo constante la intensidad de la competencia de las importaciones y eliminó, o al menos redujo considerablemente, el problema del sesgo de la variable omitida. Podría haber mantenido constante esta variable de control clave en otro nivel —como el de aquellos sectores en los que se registraba una moderada penetración de importaciones—, siempre que hubiera sido realmente constante en sus observaciones.

Sin embargo, una vez que había controlado la competencia de las importaciones, Milner tuvo que enfrentarse a otros problemas relacionados con las variables omitidas. Las dos que barajó más seriamente —a partir de una revisión de los escritos empíricos y teóricos del área— fueron: (1) que los cambios registrados en el poder de los Estados Unidos explicaban los diferentes resultados de los años veinte y setenta y (2) que la explicación se basaba en las transformaciones registradas en los procesos políticos internos del país. El intento de controlar la primera explicación ya estaba incluido en su primer diseño de investigación: como la proporción del comercio mundial en la que participaban los Estados Unidos en la década de los setenta era prácticamente igual a la de los veinte, la investigadora controló

esta dimensión del poder estadounidense en el nivel agregado de sus políticas, así como en el de los sectores industriales y empresariales. Sin embargo, no controló las diferencias que había entre el aislacionismo político de los Estados Unidos en los veinte y su posición hegemónica como líder de una alianza en los setenta; se podría profundizar más en estos factores para calibrar su potencial capacidad de sesgo.

Milner controló los procesos políticos internos comparando sectores industriales y empresas en los veinte y en los setenta, puesto que unos y otras se habían enfrentado a las mismas estructuras gubernamentales y procesos políticos. Evidentemente, su nuevo estudio de seis sectores industriales franceses que competían en el ámbito de las importaciones durante los años setenta no le ayudó a mantener constantes los procesos políticos internos, pero sí le ayudó a descubrir que el efecto causal que tenía la dependencia de las exportaciones sobre la tendencia al proteccionismo no variaba cuando lo hacían los procesos políticos internos. Al pensar cuidadosamente en las razones potenciales que había para incurrir en el sesgo de la variable omitida y diseñar su estudio para responder a esa necesidad, Milner redujo considerablemente este posible sesgo.

Sin embargo, no controló de forma explícita otras posibles variables omitidas. Su estudio se centró «en las preferencias de los grandes grupos comerciales y no examina directamente la influencia de la opinión pública, la ideología, los sindicatos, la estructura política del país ni otros posibles factores» (1988, pp. 15-16). Desde el punto de vista teórico, la decisión de no controlar estas variables omitidas se podría haber justificado alegando que no estaban relacionadas con las principales variables causales (dependencia de la exportación e inversión a escala multinacional), que, en parte, eran consecuencia de estas últimas o que no tenían efectos sobre la variable dependiente (preferencia por el proteccionismo en un nivel agregado, compuesto de empresas y sectores industriales). Sin embargo, si se podían relacionar tales variables omitidas tanto con las explicativas del estudio como con la dependiente, y, desde el punto de vista causal, eran anteriores a las explicativas, Milner tendría que haber diseñado una investigación que las hubiera controlado explícitamente¹².

Para terminar, con su procedimiento de selección de sectores industriales Milner corrió el riesgo de hacer ineficientes sus inferencias causales. Como hemos señalado, su forma de elegir los casos le permitió controlar la causa más importante de sesgo de la variable omitida mediante el mantenimiento en valores constantes de la competencia de las importaciones, variable que, desde un punto de vista teórico, se esperaba que fuera causalmente anterior a la causal clave, que se correlacionara con ella y que influyera también en las variables dependientes. Seleccionó los sectores que sufrían una mayor competencia de las importaciones y no los estratificó en función de ninguna otra variable. Posteriormente, en cada uno de los sectores de su muestra, y también en muchas empresas, estudió el grado de preferencia

por el proteccionismo (su variable dependiente) e investigó el nivel de dependencia económica del exterior (su variable explicativa).

Este procedimiento de selección resultó ineficiente para sus inferencias causales porque sus variables causales clave cambiaron menos de lo que se hubiera deseado (Milner, 1988, pp. 39-42). Aunque, en este caso, la ineficiencia no resultó ser un grave problema, sí tuvo como consecuencia que la investigadora tuviera que realizar más estudios de caso de los necesarios para llegar al mismo nivel de certidumbre en sus conclusiones (véase el apartado 2 del capítulo 6). Dicho de otro modo, si hubiera trabajado con el mismo número de casos, pero eligiéndolos de modo que, en general, variaran con la variable explicativa, podría haber extraído inferencias causales más sólidas. Es decir, que el diseño habría sido más eficiente si se hubieran seleccionado, por una parte, algunos sectores industriales y empresas que carecían de vínculos con el extranjero y, por otra, industrias y empresas en las que había un alto nivel de participación extranjera; todas ellas habían sufrido constantes apuros económicos y también la entrada de importaciones.

Los investigadores nunca pueden rechazar del todo la posibilidad de que las variables omitidas hayan sesgado sus análisis. Sin embargo, Milner, al haber intentado controlar algunas evidentes fuentes de sesgo de la variable omitida, pudo defender su hipótesis con más ahínco y de modo más convincente. Su riguroso estudio demuestra que los científicos sociales que trabajan con elementos cualitativos no tienen por qué desesperarse ante la perspectiva de extraer inferencias causales limitadas. La perfección es inalcanzable, quizá incluso indefinible, pero relacionar con cuidado la teoría y el método puede hacer que diseñemos estudios que aumenten la plausibilidad de nuestros argumentos y que reduzcan la incertidumbre de las inferencias causales.

3. Inclusión de variables irrelevantes: la ineficiencia

A causa de los problemas potenciales que presenta el sesgo de la variable omitida —que hemos descrito en el apartado anterior—, podríamos pensar ingenuamente que es esencial recoger todas las posibles variables explicativas, estimando a la vez sus efectos causales. Desde el principio hay que recordar que esto no es lo que se desprende del apartado 2 de este capítulo, en el que hemos señalado que omitir una variable explicativa que no se correlaciona con las explicativas incluidas no genera sesgo —aunque esa variable tenga un fuerte impacto causal en la dependiente— y que es un error controlar variables que son consecuencia de las explicativas. En consecuencia, *nuestro argumento no debería llevar a los investigadores a recabar información sobre todas las inferencias causales posibles o a criticar los estudios que no lo hagan.*

Por supuesto, puede darse el caso de que el investigador siga sin estar seguro de qué variables de control previas tienen un impacto causal en las incluidas o cuáles se correlacionan con éstas. En tal situación, habrá investigadores que intenten incorporar todas las variables de control cuya correlación con las variables explicativas utilizadas sea concebible, así como todas aquellas que, desde un punto de vista teórico, se pueda esperar que afecten a la variable dependiente. Es probable que el resultado de esto sea una larga lista de variables y que muchas de ellas sean irrelevantes. Este enfoque, que a primera vista parece una manera cuidadosa y prudente de evitar el sesgo de la variable omitida, de hecho nos haría correr el riesgo de elaborar un diseño de investigación cuyos resultados no podrían ser más que imprecisos. Cuando se utiliza un número relativamente escaso de observaciones, la imprecisión, tal como se analizó en el apartado 1 del capítulo 4, es un problema bastante grave para el que este «cuidadoso» diseño sería en realidad perjudicial. En el presente apartado analizamos los costes que tiene incluir variables explicativas irrelevantes y mostramos reservas fundamentales ante la pretensión de «incluirlo todo». Incorporar variables irrelevantes puede ser muy dañino. Nuestra idea principal es que aunque la variable de control no tenga ningún efecto causal en la dependiente, *cuanta mayor correlación haya entre la variable explicativa y la irrelevante de control, menos eficiente será la estimación del principal efecto causal.*

Vamos a ilustrar esta afirmación centrándonos en dos diferentes procedimientos (o «estimadores»), con el fin de calcular el efecto causal de una variable explicativa cuya inclusión sí es apropiada. La primera estimación de este efecto procede de un análisis realizado sin variables de control irrelevantes; en el segundo hay una que sí lo es. El análisis formal que figura en el recuadro posterior llega a las siguientes conclusiones respecto al valor relativo de estos dos procedimientos y al del método antes mencionado. En primer lugar, *ninguno de los estimadores está sesgado.* Esto quiere decir que, aunque se controle una variable explicativa irrelevante, el estimador habitual seguirá dándonos una respuesta media correcta. En segundo lugar, *si la variable de control irrelevante no se correlaciona con la explicativa principal, la estimación del efecto causal de la segunda no sólo estará sesgado, sino que resultará tan ineficiente como si no se hubiera incluido la variable irrelevante.* En realidad, si estas variables no están correlacionadas, se obtendrá precisamente la misma inferencia. Sin embargo, si la variable de control irrelevante está muy correlacionada con la explicativa principal, se registrará una ineficiencia considerable.

Por lo tanto, los costes de controlar variables irrelevantes son altos. Si lo hacemos, es más posible que nuestros estudios hagan estimaciones alejadas de los auténticos efectos causales. Cuando reproduzcamos un estudio, en un conjunto de datos que presenta una alta correlación entre la variable explicativa clave y una de control irrelevante incluida en el análisis, probablemente obtengamos resultados diferentes que nos lleven a otras inferencias

causales. De este modo, aunque controlemos todas las variables explicativas irrelevantes (sin cometer otros errores), obtendremos la respuesta media correcta, pero quizá en algunos de los proyectos, o en todos, nos alejemos de ella. La repetición del análisis tendrá, como promedio, el mismo efecto, pero la variable irrelevante aumentará la ineficiencia, tal como habría ocurrido si hubiéramos descartado alguna de nuestras observaciones. Las consecuencias tendrían que ser claras: al incluir una variable irrelevante ejercemos más presión sobre un conjunto de datos limitado, lo cual hace que dispongamos de menos información para cada inferencia.

Como ejemplo, pensemos de nuevo en el estudio de los golpes de Estado en los países africanos. Una investigación preliminar indicaba que el grado de represión política, la variable explicativa principal de interés, aumentaba la frecuencia de dichos golpes. Supongamos que otro investigador señalara que el primer estudio era incorrecto porque no había tenido en cuenta si cada país había logrado la independencia librándose del gobierno colonial de manera violenta o negociada. Supongamos que creemos que este investigador está equivocado y que el tipo de proceso descolonizador no influye en la variable dependiente, que es la frecuencia de los golpes (después de utilizar la principal variable explicativa, se controla la represión política). ¿Cuáles serían las consecuencias de controlar esta otra variable irrelevante?

La respuesta depende de la relación que exista entre la variable irrelevante, que evalúa el tipo de proceso descolonizador, y la explicativa principal, que mide la represión política. Si la correlación entre ambas es alta —como parece posible—, incluir estas variables de control hará que las estimaciones del efecto de la represión política sean bastante ineficientes. Para entender esto, hay que señalar que, para controlar de qué manera se accedió a la independencia, el investigador podría dividir sus categorías de regímenes represivos y no represivos teniendo en cuenta si el proceso descolonizador había sido violento o negociado. Se podría medir la frecuencia de los golpes de Estado en cada categoría para evaluar los efectos causales de la represión política, mientras que se mantiene bajo control la forma de librarse del gobierno colonial. Aunque este tipo de diseño sea una forma razonable de evitar el sesgo de la variable omitida, puede tener muchos costes: cuando la nueva variable de control no influya en la dependiente pero se correlacione con una variable explicativa incluida, se reducirá el número de observaciones de cada categoría y la estimación del principal efecto causal se realizará de forma mucho menos eficiente. Este resultado supone que se desperdicia gran parte del trabajo del investigador, ya que reducir innecesariamente la eficiencia equivale a descartar observaciones. La mejor solución es siempre recoger más observaciones, pero, si esto no es posible, los investigadores ya saben que tienen que identificar cuáles son las variables irrelevantes para no tener en cuenta sus consecuencias.

Un análisis formal de las ineficiencias de la variable incluida

Supongamos que el verdadero modelo sea $E(Y) = X_1\beta_1 + X_2\beta_2$ y $V(Y) = \sigma^2$. Sin embargo, pensamos erróneamente que la variable explicativa X_2 también ha de estar en la ecuación; de manera que hacemos la siguiente estimación:

$$E(Y) = X_1\beta_1 + X_2\beta_2 \quad (5.6)$$

sin saber que, en realidad, $\beta_2 = 0$. ¿Qué consecuencias tiene una estimación simultánea de ambos parámetros para nuestra estimación de β_1 ?

Si definimos b_1 como el estimador correcto —que se basa únicamente en la regresión entre Y y X_1 — y $\hat{\beta}_1$ como el primer coeficiente de X_1 en una regresión entre Y , por una parte, y X_1 y X_2 , por otra, es fácil mostrar que no podemos distinguir entre estos dos estimadores utilizando la falta de sesgo (una media correcta en muchos experimentos hipotéticos), ya que ambos carecen de él:

$$E(b_1) = E(\hat{\beta}_1) = \beta_1 \quad (5.7)$$

Sin embargo, los estimadores sí difieren en cuanto a su eficiencia. El estimador correcto tiene una varianza (calculada en la ecuación [3.9]) de

$$V(b_1) = \frac{\sigma^2}{\sum_{i=1}^n X_{1i}^2} \quad (5.8)$$

mientras que la del otro estimador es

$$\begin{aligned} V(\hat{\beta}_1) &= \frac{\sigma^2}{(1 - r_{12}^2) \sum_{i=1}^n X_{1i}^2} \quad (5.9) \\ &= \frac{V(b_1)}{(1 - r_{12}^2)} \end{aligned}$$

donde la correlación entre X_1 y X_2 es r_{12} (véase Goldberger, 1991, p. 245).

En la última línea de la ecuación podemos ver la relación precisa que existe entre las varianzas de los dos estimadores. Si la correlación entre las dos variables explicativas es nula, no tendrá importancia si se incluye la variable irrelevante o no, ya que ambos estimadores tendrán la misma varianza. Sin embargo, cuanto más se correlacionen dos variables, mayor será la varianza y, por tanto, menor la eficiencia de $\hat{\beta}_1$.

4. Endogeneidad

En las ciencias políticas la investigación experimental no es frecuente. No solemos tener la oportunidad de manipular las variables explicativas; sólo las observamos. Una de las consecuencias de esta falta de control es la endogeneidad, es decir, que los valores de nuestras variables explicativas a veces provienen de nuestra variable dependiente en vez de ser una de sus causas. Si hay una auténtica manipulación experimental, la dirección de la causalidad es inequívoca. Sin embargo, en muchas áreas de la investigación cualitativa y cuantitativa la endogeneidad es un problema habitual y grave¹³.

Si el investigador no controla los valores de las variables explicativas, la dirección de la causalidad siempre será un problema difícil. En la investigación no experimental —ya sea cuantitativa o cualitativa— las variables explicativas y dependientes varían por la influencia de factores que escapan al control del investigador (y que, a menudo, no se ven). Los países invaden, los oficiales traman golpes de Estado, la inflación cae, las políticas de los gobiernos se implantan, los candidatos deciden presentarse a las elecciones, los votantes eligen entre ellos y el investigador tiene que intentar elaborar un argumento para determinar causas y consecuencias.

Pensemos en un ejemplo tomado de la bibliografía sobre las elecciones al Congreso estadounidense. Muchos académicos han señalado que el drástico aumento que se registró a finales de los sesenta en las ventajas electorales de que disponían los candidatos que ya estaban en el poder se debió principalmente al incremento de los servicios que prestaban estos congresistas a su electorado. Es decir, que el franquero gratuito, los presupuestos para viajar a la propia circunscripción, el personal de que se dispone en ella para ocuparse de las peticiones concretas de los electores, la financiación de proyectos locales con fondos públicos para lograr objetivos electorales, así como otros incentivos de los congresistas, han hecho posible que éstos recaben apoyos en sus circunscripciones. Muchos ciudadanos les votan teniendo en cuenta estas cuestiones.

La hipótesis del servicio al electorado parece perfectamente razonable, pero ¿hay pruebas de ella? Numerosos investigadores han intentado presentarlas (para una revisión de esta bibliografía, véase Cain, Ferejohn y Fiorina, 1987), pero pocos son los datos que constatan tal hipótesis. El estudio modal de esta cuestión se basa en medidas del servicio que ha dado a su circunscripción cada uno de los congresistas de una muestra y en las que indican la proporción de votos que recibe esa clase de candidatos. Posteriormente, utilizando un análisis de regresión, los investigadores hacen una estimación del impacto causal que tiene ese servicio sobre el voto. Resulta sorprendente que muchas de esas estimaciones indiquen que el efecto es nulo o incluso negativo.

Es probable que el problema de la endogeneidad explique esos resultados paradójicos. Dicho de otro modo, los congresistas que corran un mayor

análisis no corregido, el efecto del servicio presentará un límite más bajo, lo cual hará que resulte más plausible la hipótesis del servicio al electorado.

De este manera, aunque no podamos evitar el sesgo de endogeneidad, a veces sí podremos mejorar las inferencias a posteriori haciendo una estimación del nivel de sesgo. Como mínimo, esto nos permitirá determinar su dirección y quizá nos indique el límite superior o inferior de la estimación correcta. En el mejor de los casos, podremos utilizar esta técnica para hacer inferencias que no tengan sesgo alguno.

4.2 Precisar cuáles son los componentes de la variable dependiente

Una de las formas de evitar el sesgo de endogeneidad es redefinir la variable dependiente para dotarla de sus propios componentes dependiente y explicativo. Este último obstaculiza nuestro análisis a través de un mecanismo de retroalimentación, es decir, influye en la variable causal (explicativa) clave. La otra parte de nuestra variable dependiente es realmente dependiente: es una función, y no una causa, de la variable explicativa. El objetivo de esta forma de evitar el sesgo de endogeneidad es identificar y medir sólo la parte dependiente de nuestra variable dependiente.

Por ejemplo, King (1991a), al estudiar la hipótesis del servicio al electorado, desgajó del total de votos recibido por un congresista estadounidense la porción que se debía únicamente al hecho de que ya estuviera en el poder. En los últimos años, la ventaja electoral de los que ya son congresistas se ha cifrado entre un 8 y un 10% de los votos; proporción que hay que comparar con el 52% aproximado que constituye la base electoral de la mayoría de los que, en este sistema bipartidista, ya ocupan un escaño en el Congreso. Posteriormente, King calculó, mediante un procedimiento estadístico, la ventaja que suponía ser congresista —elemento que era el único dependiente de la variable dependiente— y utilizó esta cifra en lugar del voto bruto para hacer una estimación de los efectos que tenía el servicio al electorado. Como la ventaja de los congresistas en número de votos, al representar sólo una pequeña porción del total, no iba a tener un gran impacto en la propensión de ese tipo de candidatos a realizar servicios para el electorado, King evitó el sesgo de endogeneidad. Sus resultados señalaban que, si se añadían 10.000 dólares al presupuesto que tenía el legislador estatal medio para (entre otras cosas) dar servicios al electorado, se estaba concediendo a este candidato una ventaja de 1,54 puntos (más o menos 0,4%) en las elecciones siguientes. De esta manera, el autor estaba dando la primera base empírica a la hipótesis del servicio al electorado.

4.3 Convertir la endogeneidad en un problema de variable omitida

Siempre se puede abordar la endogeneidad como si fuera un sesgo de la variable omitida, tal como demostramos a continuación con el famoso ejemplo del estudio sobre sistemas electorales comparados. Para las primeras generaciones de politólogos, uno de los grandes enigmas del análisis político era la caída de la República de Weimar y su sustitución por el régimen nazi a principios de los años treinta. Una de las explicaciones, que se basaba en algunos rigurosos y convincentes estudios de caso de la Alemania de esa época, señalaba que la causa principal había sido la consagración en la constitución de esa república de un sistema de representación proporcional para las elecciones. En pocas palabras, la idea es que el sistema proporcional permite que estén en el parlamento los pequeños partidos que representan a determinadas ideologías, intereses o grupos religiosos. En este tipo de ordenamiento electoral los candidatos no se ven en la necesidad de negociar sus posturas para tener éxito en las elecciones, como sí ocurre en aquellos sistemas en los que sólo hay un ganador por circunscripción. En consecuencia, el parlamento se llena de pequeños grupos ideológicos que ni quieren ni pueden trabajar juntos. La parálisis y la frustración resultantes harán posible que uno de estos grupos —en este caso los nacionalsocialistas— tome el poder (para la exposición clásica de esta teoría, véase Hermens, 1941).

El argumento del párrafo anterior se fue elaborando en varios importantes estudios de caso que se ocuparon de la caída de la República de Weimar. Hubo historiadores y politólogos que achacaron este hecho al éxito electoral de los pequeños partidos y a su falta de voluntad negociadora en el Reichstag. Este planteamiento tiene muchos problemas, como siempre ocurre cuando se explica un resultado complejo a partir de un único factor, pero sólo vamos a centrarnos en el de la endogeneidad. Detrás de esta explicación subyace un mecanismo causal que se encadena de la siguiente manera: se implantó la representación proporcional y esto hizo posible que partidos pequeños que contaban con una escasa base electoral consiguieran escaños en el Reichstag (y en éstos se incluían aquellos que, como los nacionalsocialistas, querían acabar con dicha institución). El resultado fue que el funcionamiento del Reichstag alcanzó un punto muerto y la frustración se extendió entre las masas. A su vez, esto condujo a un golpe de Estado propiciado por uno de los partidos.

Sin embargo, otros estudios —tanto de Alemania como de otras consecuencias observables— señalaron que la fragmentación de los partidos no sólo era consecuencia de la representación proporcional. Algunos investigadores argumentaron que si dicha fragmentación condujo a la implantación de la representación proporcional, también podía haber sido su causa. Al aplicar la misma variable explicativa a otras observaciones (en consonancia con la regla que indicamos en el capítulo 1: intentar probar las hipó-

tesis con datos diferentes a los que las produjeron), los investigadores se dieron cuenta de que en aquellas sociedades en las que muchos grupos de mentalidad estrecha y arraigada se oponen a otros minoritarios, étnicos o religiosos, es más probable que se opte por la representación proporcional, ya que éste es el único sistema electoral que ambas facciones pueden aceptar. Una mirada más atenta a la situación de la política alemana antes de que se introdujera este tipo de representación confirmó esta idea, ya que se localizaron muchas facciones pequeñas. La representación proporcional no las había creado, aunque puede que sí facilitara la expresión de sus ideas en el parlamento. Tampoco las facciones eran la única causa de la representación proporcional; sin embargo, sí parecía que tanto la adopción del sistema proporcional como la gran división del parlamento eran consecuencia de la fragmentación social (véase Lakeman y Lambert, 1955, p. 155, para una de las primeras explicaciones de este argumento).

De este modo, hemos transformado un problema de endogeneidad en un sesgo de la variable omitida. Esto quiere decir que la fragmentación social previa es una variable omitida que lleva a la representación proporcional; desde el punto de vista causal es anterior a ésta y, en parte, condujo a la caída de Weimar. Al transformar el problema de esta manera, los académicos consiguieron abordar mejor el problema, ya que podían medir de manera explícita esta variable omitida y controlarla en estudios posteriores. En este ejemplo, una vez que se incluyó y controló dicha variable omitida, los investigadores se dieron cuenta de que existía una posibilidad razonable de que la aparente relación causal que había entre la representación proporcional y la caída de la República de Weimar fuera casi completamente falsa.

La relación que existe entre los sistemas electorales y la democracia sigue suscitando polémica, aunque el estudio de este tema ha avanzado mucho desde los primeros trabajos que se hicieron. Los investigadores han ampliado esta área, dejando a un lado los intensos estudios de caso, que apenas se preocupaban de la lógica de las explicaciones, para hacer trabajos que se basan en muchas observaciones de determinadas consecuencias y así resolver gradualmente algunos aspectos relacionados con la medida y, finalmente, con la inferencia. Con este cambio, han logrado separar de modo más sistemático los efectos exógenos de los endógenos.

4.4 Seleccionar observaciones para evitar la endogeneidad

La endogeneidad es un problema muy común en gran parte de las obras que se ocupan del impacto que tienen las ideas en las políticas (Hall, 1989; Goldstein y Keohane, 1993). En la medida en que las ideas *reflejan las condiciones* en las que operan los actores políticos —por ejemplo, las circunstancias materiales que generan sus propios intereses en ese sentido—, el análisis del impacto que tienen sobre las políticas siempre será suscepti-

ble de incurrir en un sesgo de la variable omitida: las ideas de los actores se correlacionan con una variable omitida anterior en sentido causal —los intereses materiales— que influye en la variable dependiente —la estrategia política (véase el apartado anterior). Además, en la medida en que las ideas funcionan como *racionalizaciones* de políticas que se han llevado a cabo en otros ámbitos, también pueden ser simples *consecuencias* de esas políticas, más que sus causas. En tales circunstancias, las ideas son endógenas: puede parecer que explican las estrategias de los actores, pero, en realidad, son su resultado.

Desde el punto de vista metodológico, lo más difícil a la hora de estudiar el impacto de las ideas sobre las políticas es compensar la influencia que tienen, en un determinado problema de investigación, el problema de la endogeneidad y el de la variable omitida, que tan relacionados están. Para poder señalar la importancia causal de las ideas, hay que demostrar que algunas ideas —o un aspecto de las mismas— que mantienen los que elaboran las políticas influyen en aquellas que realmente se llevan a cabo, y que no sólo reflejan esas mismas políticas o los previos intereses materiales de sus impulsores. Los investigadores de esta área tienen que ser especialmente cuidadosos a la hora de definir el efecto causal del interés. En concreto, tienen que comparar la variable dependiente observada (las políticas) y la explicativa (las ideas de los individuos) con una situación contrafáctica —definida con precisión— en la que la variable explicativa adopte un valor diferente: cuando los individuos relevantes tienen ideas diferentes.

El análisis comparado es una buena forma de determinar si un conjunto de ideas concreto es exógeno o endógeno. Por ejemplo, en un reciente estudio acerca del papel de las ideas en la adopción de políticas económicas estalinistas en países socialistas, Nina Halpern (1993) hace este tipo de análisis. Su hipótesis es que la doctrina de la planificación estalinista, en la que creían los mandatarios de Europa del Este y de China, ayuda a explicar las políticas económicas que implantaron cuando llegaron al poder en estos países después de la Segunda Guerra Mundial. Esta hipótesis concuerda con el hecho de que tales líderes tuvieron ideas estalinistas y aplicaron políticas de ese tipo, pero una simple correlación no demuestra que haya causalidad. En realidad, puede que la endogeneidad esté funcionando: las políticas estalinistas podrían haber generado ideas que las justificaran o prever que habría que seguir tales políticas quizá produjera esas ideas.

Aunque Halpern no utiliza este lenguaje, sí recurre a un procedimiento similar al que se ha analizado en el apartado anterior: transformar la endogeneidad en sesgo de la variable omitida. La principal hipótesis alternativa que esta autora considera es que Europa del Este y los estados comunistas de Asia sólo implantaron economías centralizadas después de la Segunda Guerra Mundial por el poderío militar y la influencia política de la Unión Soviética. El contrafáctico que plantea esta hipótesis es que, aunque los europeos orientales y los chinos no hubieran creído en la conveniencia de las

economías planificadas que dictaba el estalinismo, de todas maneras se habrían implantado esa clase de economías y habrían aparecido ideas que las justificaran.

Posteriormente, Halpern señala que en los países de Europa del Este que fueron ocupados por el Ejército Rojo, la adopción de una economía centralizada podría explicarse más por el poderío soviético que por la creencia en la superioridad de las doctrinas estalinistas: «... la explicación alternativa de que la elección no fue más que una respuesta a las órdenes de Stalin es imposible de refutar» (1993, p. 89). En consecuencia, Halpern busca posibles observaciones en las que no aparezca esta fuente de sesgo de la variable omitida y encuentra las políticas que se implantaron en China y Yugoslavia, los dos países socialistas más grandes de los no ocupados por tropas soviéticas después de la Segunda Guerra Mundial. Como China era un país enorme que había tenido su propia revolución, Stalin no podía indicarle qué políticas tenía que practicar. Los comunistas yugoslavos también habían llegado al poder sin la ayuda del Ejército Rojo, y el mariscal Tito demostró su independencia respecto a las órdenes de Moscú desde el final de la guerra.

China implantó una economía centralizada sin estar bajo la dominación política o militar de la Unión Soviética y en Yugoslavia las medidas de tipo estalinista se adoptaron a pesar de las políticas soviéticas. De estos datos, Halpern infiere que, en estos casos, el poder soviético no explica por sí solo los cambios en las políticas. Además, por lo que a China se refiere, también considera, y rechaza, otra hipótesis alternativa según la cual las ideas serían endógenas: situaciones económicas parecidas habrían aconsejado transplantar los métodos de planificación estalinista a China.

Halpern, una vez que se planteó y rechazó hipótesis alternativas que mantenían ideas endógenas bien en relación al poder soviético o a las condiciones económicas, pudo indicar que la adopción por parte de China (y, hasta cierto punto y durante menos tiempo, de Yugoslavia) de la doctrina estalinista aportaba bases para el acuerdo y resolvía la incertidumbre de estos regímenes postrevolucionarios. Aunque este tipo de análisis sigue siendo bastante provisional, a causa del escaso número de observaciones de su propia teoría que contempló la autora, sí aporta razones para creer que, en la situación estudiada, las ideas no eran del todo endógenas y tuvieron una función causal.

Este ejemplo pone de manifiesto de qué manera podemos, en primer lugar, trasladar una preocupación general sobre la endogeneidad a posibles fuentes específicas de sesgo de la variable omitida, para buscar después un subconjunto de observaciones en el que podrían no estar presentes esas causas de sesgo. En este caso, al transformar el problema en una cuestión de sesgo de la variable omitida, Halpern pudo comparar otras hipótesis explicativas con resultados especialmente productivos para su hipótesis principal. Se planteó diversas hipótesis explicativas alternativas para explicar la adopción de economías planificadas y se dio cuenta de que sólo en China, y hasta cierto punto en Yugoslavia, era razonable pensar que las doctrinas

estalinistas (las ideas en cuestión) fueran en gran medida exógenas. En consecuencia, centró su atención en China y Yugoslavia. Si no hubiera diseñado su estudio con cuidado para poder enfrentarse al problema de la endogeneidad, sus conclusiones habrían sido mucho menos convincentes; pensemos, por ejemplo, que hubiera intentado demostrar sus planteamientos utilizando Polonia y Bulgaria!

4.5 Precisar cuáles son los componentes de la variable explicativa

En este apartado presentamos el quinto y último método para eliminar el sesgo que produce la endogeneidad. Su objetivo es dividir una variable explicativa potencialmente endógena en dos componentes: uno claramente exógeno y otro que sea, al menos en parte, endógeno. Posteriormente, el investigador sólo utilizará en su análisis causal la parte exógena de la variable explicativa.

En el estudio realizado por Verba, Schlozman, Brady y Nie (1993) sobre participación política voluntaria, encontramos un ejemplo de este método. A los autores les interesaba explicar por qué los afroamericanos son mucho más activos políticamente que los latinos, teniendo en cuenta que ambos grupos están igualmente desfavorecidos. Se dieron cuenta de que hay diversos factores que contribuyen a explicar la diferencia, entre ellos el poco tiempo transcurrido desde que se ha emigrado a los Estados Unidos y el conocimiento del idioma. Una de sus principales variables explicativas era la asistencia a servicios religiosos (en iglesias o sinagogas, por ejemplo). Evidentemente, los investigadores no podían controlar si los individuos asistían a esas reuniones; de ahí que no pudiera eliminarse la posibilidad de endogeneidad. En realidad, sospechaban que algunos latinos, y en mayor medida los afroamericanos, asistían a servicios religiosos porque eran políticamente activos. Quizá los que tenían interés en política entraban a formar parte de una iglesia porque ésta les daba la oportunidad de aprender o porque estaba muy politizada. Un clero muy consciente de su labor política podía adiestrar a los feligreses para ese tipo de actividad o estimularles políticamente. Dicho de otro modo, la orientación de la causalidad podía ir desde la política a las experiencias no políticas y no al revés.

Verba y sus colegas solucionaron este problema analizando los componentes de su principal variable explicativa. Para ello, partieron de la base de que las instituciones religiosas influyen en la participación política de dos maneras. En primer lugar, en ellas los individuos aprenden habilidades sociales (como dar un discurso o dirigir una reunión). A su vez, ese aprendizaje hace que el ciudadano sea más competente a la hora de participar en la vida política y que esté más dispuesto a hacerlo. En segundo lugar, los ciudadanos reciben constantes estímulos políticos (por ejemplo, a través de discusiones o cuando otros miembros del grupo les piden que se impliquen

directamente), y esta situación también tiene que influir en su actividad política. Estos autores señalaron que el primer componente es principalmente exógeno, mientras que el segundo es, al menos en parte, exógeno: o sea, que procede parcialmente del grado de participación política de los individuos (la variable dependiente).

Posteriormente, los autores llevaron a cabo un estudio auxiliar para contrastar su hipótesis en relación a los componentes exógeno y endógeno de la participación en los servicios religiosos. Desde el principio reconocieron que la posibilidad de que el individuo adquiera habilidades sociales en su iglesia depende de la estructura organizativa de ésta. Las que tengan una estructura jerárquica, donde el clero sea elegido por órganos rectores centralizados y los feligreses apenas puedan opinar sobre el funcionamiento interno, ofrecerán menos oportunidades de aprender habilidades sociales que aquellas cuya base cooperativa conceda a los miembros un importante papel en el gobierno de la organización. La mayoría de los afroamericanos pertenecen a iglesias protestantes que funcionan de manera muy participativa, mientras que los latinos son miembros de congregaciones católicas organizadas de modo jerárquico. Los autores mostraron que esta diferencia es la que explica la probabilidad de que adquieran habilidades sociales. Por ejemplo, señalaron que en estos dos grupos, así como en los estadounidenses blancos, el factor que influye en el aprendizaje de esas habilidades es el tipo de confesión religiosa y no los componentes étnicos, ni otros rasgos sociales ni, especialmente, la participación política.

Después de convencerse de que la adquisición de habilidades sociales era realmente exógena a la participación política, Verba y los otros autores calibraron este factor en los servicios religiosos y lo utilizaron como variable explicativa, en vez de usar para tal fin la asistencia a tales servicios. Este procedimiento solucionó el problema de la endogeneidad, ya que ahora se habían analizado los componentes de la variable explicativa para poder incluir únicamente los exógenos.

Este estudio auxiliar aportó más pruebas que demostraban la resolución del problema de la endogeneidad, ya que la pertenencia de los latinos y afroamericanos a las iglesias no podía explicarse de manera coherente en función de su participación política concreta, puesto que su vinculación con ellas comienza desde la infancia a través de la familia. Las razones que explican el hecho de que los afroamericanos sean en su mayoría protestantes hay que buscarlas en la historia de la esclavitud en los Estados Unidos y en la de las organizaciones que se formaron en las plantaciones del sur, mientras que el catolicismo de los latinos procede de la conquista española de Hispanoamérica. Tampoco puede atribuirse la diferente estructura institucional de las congregaciones católicas y de las protestantes al interés que tienen los dirigentes de cada una de ellas por participar en la política estadounidense actual. En realidad, hay que retrotraerse a la época de la Reforma protestante para poder explicar tal diferencia.

5. Asignación de valores a la variable explicativa

En el apartado cuarto del capítulo 4 señalamos que los experimentos mejor dirigidos tienen dos ventajas: el control de la selección de observaciones y de los valores que se asignan a las variables explicativas en las unidades. Sólo se analizó la selección en ese punto. Como en este capítulo ya nos hemos ocupado del sesgo de la variable omitida y de otros problemas metodológicos, podemos enfrentarnos ahora al control de la asignación de valores.

En un experimento médico, los tratamientos son un fármaco que se está probando y un placebo, los cuales se administran aleatoriamente a los pacientes. Aquí se da prácticamente la misma situación que aparece con una selección aleatoria de observaciones: la asignación aleatoria es muy útil cuando existe una gran número de observaciones, pero no suele ser la estrategia óptima si n es pequeño. Cuando es grande, asignar valores aleatorios a las variables explicativas acaba con la posible endogeneidad (porque dichas variables no pueden verse influidas por la variable dependiente) y con el error de medida (siempre que se registre exactamente qué tratamiento se administra). Quizá lo más importante sea que la asignación de valores en estudios con un n grande hace extremadamente improbable la aparición de sesgo de la variable omitida, porque una variable explicativa cuyos valores se hayan asignado aleatoriamente no se correlacionará con ninguna variable omitida, ni siquiera con las que influyan en la dependiente. Por lo tanto, en estudios de n grande, la asignación aleatoria de valores hará que las variables omitidas no sean dañinas, que no produzcan sesgo. Sin embargo, si se dispone de un número reducido de observaciones, es muy fácil que una variable asignada de forma aleatoria se correlacione con alguna variable omitida relevante y que tal correlación produzca sesgo de la variable omitida. De hecho, el ejemplo del sesgo de selección mostraba cómo una variable asignada aleatoriamente se correlacionaba con una variable dependiente observada; es muy fácil que ocurra exactamente lo mismo si el número de observaciones es pequeño.

Aunque al realizar experimentos muchas veces se pueden fijar los valores de las variables explicativas, los investigadores cualitativos no suelen tener esa suerte. Cuando el objeto de estudio determina el valor de sus propias variables explicativas, aumentan considerablemente las posibilidades de sesgo de selección, de endogeneidad y las de otros tipos de sesgo y de ineficiencia. Por ejemplo, si en un experimento se estuviera estudiando la eficacia política que tiene participar en manifestaciones, se elegiría aleatoriamente a ciertas personas para que acudieran a una manifestación y a otras para que se quedaran en casa, con el fin de calibrar después las diferencias que se registran entre los dos grupos del experimento en relación a la eficacia (o quizá para compararlos en función del cambio que registra ese indicador en una medida realizada antes del experimento y después de él). Sin embargo, en el trabajo no experimental, es frecuente que los pro-

Un análisis formal de la endogeneidad

Este modelo formal demuestra el sesgo que aparece cuando un diseño de investigación padece endogeneidad y no se hace nada por evitarlo. Supongamos que tenemos una variable explicativa X y una dependiente Y . Nos interesa el efecto causal que tiene X sobre Y , y utilizamos la siguiente ecuación:

$$E(Y) = X\beta \quad (5.10)$$

Ésta también puede expresarse como $Y = X\beta + \varepsilon$, en la que $\varepsilon = Y - E(Y)$ se denomina el «error» o «término perturbador». Supongamos, además, que hay endogeneidad, o sea, que X también depende de Y :

$$E(X) = Y\gamma \quad (5.11)$$

¿Qué ocurre si no prestamos atención a la parte recíproca de la relación expresada en la ecuación (5.11) y hacemos una estimación de β como si sólo fuera cierta la ecuación (5.10)? Dicho de otro modo, estimamos β (presuponiendo incorrectamente que $\gamma = 0$) mediante la ecuación habitual:

$$b = \frac{\sum_{i=1}^n X_i Y_i}{\sum_{i=1}^n X_i^2} \quad (5.12)$$

Para evaluar este estimador utilizamos la propiedad de falta de sesgo y, por lo tanto, calculamos su valor esperado:

$$E(b) = E\left(\frac{\sum_{i=1}^n X_i Y_i}{\sum_{i=1}^n X_i^2}\right) \quad (5.13)$$

pios objetos de estudio decidan si quieren participar. En estas circunstancias, hay otros rasgos individuales (como el hecho de que los individuos sean jóvenes o no, que sean estudiantes y, si lo son, la cercanía del campus al lugar de la manifestación, etc.) que influirán en la decisión de manifestarse. Además, es evidente que muchos de estos factores también pueden correlacionarse con la variable dependiente, que es la eficacia política.

Pensemos en otro ejemplo en el que las unidades de análisis sean más grandes y menos frecuentes: el clásico problema del impacto que tiene la acumulación de armamento en la posibilidad de guerra. ¿El tamaño del presupuesto nacional para esta partida aumenta las probabilidades de que el país participe en una guerra? La variable explicativa es el presupuesto para

$$\begin{aligned} &= E\left(\frac{\sum_{i=1}^n X_i(X_i\beta + \varepsilon_i)}{\sum_{i=1}^n X_i^2}\right) \\ &= \beta + \frac{\sum_{i=1}^n C(X_i, \varepsilon_i)}{\sum_{i=1}^n V(X_i)} \\ &= \beta + \text{sesgo} \end{aligned}$$

donde el sesgo = $\sum_{i=1}^n C(X_i, \varepsilon_i) / \sum_{i=1}^n V(X_i)$. Normalmente, la covarianza de X_i y el término perturbador ε_i , $C(X_i, \varepsilon_i)$ valen cero, de manera que el término de sesgo también tendrá este valor. En consecuencia, el valor esperado de b es β y, por tanto, carece de sesgo. Suele ocurrir que, después de que se tiene en cuenta X para predecir Y , la parte que queda (ε) no se correlaciona con X . Sin embargo, en el presente ejemplo, después de que se contempla el efecto de X , sigue quedando alguna variación por la retroalimentación que procede del efecto causal de Y sobre X . De este modo, la endogeneidad implica que, en general, el segundo término de la última línea de la ecuación (5.13) no valdrá cero y que la estimación estará sesgada.

La dirección del sesgo depende de la covarianza, ya que la varianza de X siempre es positiva. Sin embargo, en los pocos casos en los que ésta sea extremadamente grande, amollará a la covarianza y hará que el término de sesgo resulte insignificante. En el texto se ofrece un ejemplo con una buena interpretación de este término de sesgo.

armamento (quizá el porcentaje en relación al PNB o los cambios que ha sufrido el propio presupuesto); la variable dependiente es la presencia o ausencia de guerra en un período determinado posterior al momento en que se ha medido la variable explicativa. En un diseño de investigación ideal sería el investigador el que asignara valores a la variable explicativa, eligiendo diversos países para su estudio y determinando el presupuesto armamentístico de cada gobierno (mediante valores aleatorios o utilizando quizá una de las variables «intencionadas» que analizamos más adelante). Es evidente que esto no es factible. Lo que se hace en realidad es medir los valores de la variable explicativa (el tamaño del presupuesto armamentístico) que cada gobierno se asigna. Por supuesto, el problema es que estos valores autoa-

signados de la variable explicativa no son ajenos a la dependiente —la probabilidad de ir a la guerra—, como lo hubieran sido si hubiéramos podido elegirlos nosotros. En este caso, hay un claro problema de endogeneidad: el valor de la variable explicativa está influido por los valores que se cree tendrá la dependiente: la amenaza de guerra percibida. La endogeneidad también plantea problemas para estudiar la relación causal entre las alianzas y la guerra. Las naciones eligen sus alianzas; los investigadores, sin adjudicar países a esas alianzas, estudian el impacto que tienen sobre las guerras. Por lo tanto, las alianzas no tendrían que considerarse variables explicativas exógenas en los estudios bélicos, ya que suelen formarse antes que la propia guerra.

Estos ejemplos demuestran que la endogeneidad no siempre es un problema que haya que precisar, sino que a veces constituye una parte integral del proceso por el que el mundo genera observaciones. Con frecuencia, resulta bastante difícil averiguar cómo se han determinado los valores de las variables explicativas, y no suele ser posible recurrir a ningún procedimiento automático para resolver los problemas que plantea. No obstante, al investigar no puede eludirse esta tarea.

Como la probabilidad de que, en cualquier prueba de un experimento hipotético, el uso de criterios aleatorios de selección o de asignación de valores produzca sesgo disminuye rápidamente a medida que aumenta el número de observaciones, es útil recurrir a tales procedimientos aunque no se disponga de un gran número de unidades. Si la cantidad de unidades es «suficientemente grande» —lo cual definimos con precisión en el apartado 2 del capítulo 6—, la selección aleatoria de unidades cumplirá automáticamente el presupuesto de independencia condicional del apartado tercero del capítulo 3. Sin embargo, cuando sólo existan unos pocos ejemplos del fenómeno que nos interesa o únicamente podamos recabar información acerca de un pequeño número de observaciones, como ocurre con frecuencia en la investigación cualitativa, la respuesta no estará en la selección y asignación de valores con procedimientos aleatorios. Ni siquiera los experimentos controlados, si son posibles, serán una solución cuando no se disponga de un número suficiente de observaciones.

Los cualitativistas, al enfrentarse a estos problemas, tienen que preguntarse si es posible aumentar el número de observaciones que van a investigar, ya que, a falta de reunir las todas, *el procedimiento más fiable es recoger de forma aleatoria datos relacionados con un gran número de unidades y aleatorizar la asignación de valores para las variables explicativas. Sin embargo, si esto no es posible, no hay que seleccionar las observaciones aleatoriamente, sino que es preciso utilizar el conocimiento que ya tenemos de las observaciones disponibles* —que se basa en investigaciones anteriores, suposiciones fundadas o valoraciones de otros expertos en el área— y *hacer la selección de las observaciones y (si se puede) la asignación de valores a las variables explicativas de tal modo que se evite el sesgo y las ineficiencias. Si*

el sesgo es inevitable, al menos habrá que intentar averiguar su dirección y su posible magnitud. Si todo esto falla —es decir, si sabemos que hay sesgo pero no podemos determinar su dirección o envergadura—, nuestra investigación mejorará si, por lo menos, describimos nuestros resultados con un mayor grado de incertidumbre. Si comprendemos los problemas inferenciales que aborda este libro, estaremos más preparados que cualquier generador de números aleatorios para enfrentarnos a todos estos dilemas. En cualquier caso, *en todos los estudios tendría que haber un apartado o capítulo en el que se explicaran cuidadosamente los procesos de selección y de asignación de valores.* Habría que mencionar en él qué reglas se han utilizado, incluir un listado detallado de todas las previsibles fuentes de sesgo ocultas y señalar si se ha hecho algo para combatirlas.

6. Controlar la situación que se investiga

Si se seleccionan de forma deliberada observaciones sin tener en cuenta las variables de control relevantes y tampoco otros problemas inferenciales, no se cumplirá el presupuesto de homogeneidad de las unidades. Tenemos que asegurarnos de que los valores de la variable explicativa en las observaciones elegidas se midan con el mínimo error posible, de que no se correlacionen con alguna variable explicativa omitida clave y de que no estén parcialmente determinados por la dependiente. Es decir, hay que calibrar de manera adecuada problemas como el error de medida, las variables omitidas y la endogeneidad, que ya se han analizado en este capítulo. Si éstos persisten, a pesar de todos nuestros esfuerzos, por lo menos tendremos que reconocerlos, sopesarlos e intentar corregirlos.

En estudios que tienen un n pequeño los controles son de por sí difíciles de diseñar, pero suele ser absolutamente esencial ocuparse de ellos para evitar el sesgo. Por desgracia, hay muchos investigadores cualitativos que apenas utilizan controles o que no lo hacen en absoluto. Por ejemplo, Bollen, Entwisle y Alderson (1993) han señalado, en un estudio sobre libros y artículos sociológicos, que alrededor de un cuarto de los investigadores no utilizan ningún método de control.

Supongamos, por ejemplo, que nos interesara el efecto causal que tiene un año de cárcel en el grado de radicalismo político que muestran las personas. Lo ideal sería un estudio auténticamente experimental en el que se utilizara una muestra aleatoria de un gran número de ciudadanos y que a la mitad de ellos se les adjudicara aleatoriamente un año de prisión, para medir después la radicalidad de las ideas políticas de uno y otro grupo. El efecto causal estimado sería la diferencia media entre las posturas de los dos grupos al final del año. Con un n grande, cabría presuponer la existencia de independencia condicional y probablemente la inferencia causal fuera sensata. No hace falta decir que tal estudio es imposible.

Sin embargo, para continuar con nuestro argumento, supongamos que se llevara a cabo el experimento, pero sólo con pocas personas. Teniendo en cuenta los problemas que se han analizado en el segundo apartado del capítulo 4, esta escasa muestra, aunque se hubiera seleccionado aleatoriamente y la asignación de valores también se hubiera hecho con este método, es probable que no cumpliera la independencia condicional y, por lo tanto, sería necesario aplicar algún control explícito. Una sencilla forma de hacerlo sería medir las ideas políticas radicales antes del experimento. Después, nuestra estimación causal sería la diferencia que hubiera entre los dos grupos en relación al *cambio* operado en las ideas políticas radicales. Este procedimiento nos permitiría controlar una situación en la que ambos grupos no son idénticos respecto a esta única variable antes de realizarse el experimento. Para comprender cómo hay que calcular el efecto causal en este contexto conviene recordar el problema fundamental de la inferencia causal. Lo ideal sería que tomáramos a un solo individuo, esperaríamos durante un año controlando cuidadosamente su medio para mantenerlo constante, excepto en lo relativo al paso del tiempo y a los acontecimientos del mundo exterior, y que después midiéramos la radicalidad de sus ideas políticas. Simultáneamente, tomaríamos a ese individuo en el mismo momento y lo meteríamos en la cárcel durante un año, para medir después el mismo indicador. La diferencia entre estas dos medidas es la definición del efecto causal que tiene la cárcel sobre las ideas políticas de esta persona¹⁴. El problema fundamental consiste que sólo podemos observar las ideas de esta persona en una de las situaciones. Es evidente que el mismo individuo no puede estar, a la vez, dentro y fuera de la cárcel.

Mediante el control se intenta evitar el problema fundamental de la manera más directa. Como no podemos observar las ideas de esta persona en ambas situaciones, buscamos dos individuos (o, más probablemente, dos grupos de individuos) que sean lo más parecidos posible, excepto en lo tocante a la variable explicativa (que hayan ido o no a la cárcel). Tampoco hacemos la selección teniendo en cuenta el grado de radicalidad de las personas. En principio, podríamos seleccionar una muestra de individuos que hubieran sido puestos en libertad recientemente y buscar después personas *equiparables* a cada ex prisionero (lo más parecido posible a éste, pero sin haber estado en la cárcel). Quizá, en primer lugar, podríamos entrevistar a alguien que hubiera salido de la cárcel y, partiendo del conocimiento de su historia y características, ir en busca de personas equiparables (que tuvieran un perfil demográfico similar y que quizá fueran del mismo barrio y hubieran ido al mismo colegio).

Por definición, las variables con las que equiparamos a los individuos son las mismas en todos los grupos y las controlaremos cuando calculemos el efecto causal de la excarcelación. Ejercer controles es un proceso difícil porque hay que aplicarlos a todas las variables que puedan ser confusas. Si no equiparamos en función de una determinada variable, si no podemos

controlarla de ninguna otra manera e influye en la variable dependiente a la vez que se correlaciona con la explicativa (afecta a la radicalidad de las ideas y no es la misma en los prisioneros y en los que no lo son), la estimación de nuestro efecto causal estará sesgada.

Cuando se comparan países desde el punto de vista político, ejercer controles para lograr la homogeneidad de las unidades es difícil: cualquier pareja de países diferirá en innumerables dimensiones. Por ejemplo, Bélgica y Holanda podrían parecerle a un observador poco preparado países «muy parecidos», en el sentido que señalan Przeworski y Teune (1982): son pequeñas democracias europeas con economías abiertas y sus vecinos no les amenazan. Por lo tanto, en muchos sentidos, sería factible compararlos (Katzenstein, 1985). Sin embargo, son muy diferentes desde el punto de vista lingüístico o religioso y en cuanto a sus recursos fundamentales, al inicio de su industrialización y a muchos otros factores políticos relevantes. En consecuencia, cualquier diseño de investigación comparado que pretenda estudiar el conjunto de la vida política en estos dos estados y que sólo se centre en ellos correrá el riesgo de ser impreciso.

Si lo que pretendemos es hacer una comparación general entre Bélgica y Holanda, se podrá evitar tal imprecisión. Sin embargo, supongamos que el investigador tenga un objetivo más específico: estudiar el impacto de ser una potencia colonial sobre las estrategias políticas que siguen los gobiernos de las pequeñas democracias europeas. En ese caso, sería posible comparar las políticas de Bélgica, Holanda y Portugal con las de pequeños estados no coloniales como Austria, Suecia, Suiza y Noruega. Éste podría ser un diseño de investigación valioso, pero no estaría controlando innumerables factores —aparte de la historia colonial— que hacen que estos países sean diferentes. Un investigador sensible al problema de la homogeneidad de las unidades podría pensar en otro diseño —quizá alternativo, aunque sería mejor que fuera un complemento para el anterior— en el que estudiará las políticas belgas, holandesas y portuguesas antes y después de la pérdida de sus respectivas colonias. En este diseño, Bélgica no sería una «única observación», sino el escenario de un análisis controlado, antes y después de que se concediera la independencia a sus colonias a principios de los años sesenta. Muchos de los factores que diferencian a Bélgica de Portugal y de Holanda —y no digamos de los países que carecen de un pasado colonial— se controlarían automáticamente en este diseño de series temporales. En realidad, ambas comparaciones —entre países y entre diversos momentos de cada país— presentan problemas en cuanto a la homogeneidad de las unidades. Las diversas naciones difieren en muchas formas no controladas y no medidas, que podrían ser relevantes para el tema de la investigación, y una nación que sea observada en diferentes momentos también presenta estas disparidades. Sin embargo, las *diferencias* no serán las mismas. Ninguna comparación (ya sea en el tiempo o en el espacio) constituye, ni mucho menos, un experimento perfectamente controlado, pero si juntamos ambos

enfoques podremos lograr más pruebas de nuestra hipótesis que si sólo utilizamos uno de ellos.

La estrategia de la selección intencionada conlleva algunos peligros ocultos de los que los investigadores tendrían que ser conscientes, especialmente al intentar equiparar observaciones con el fin de controlar variables potencialmente relevantes. El principal peligro se encuentra en un tipo de sesgo de la variable omitida especialmente engañoso. Imaginémoslo el siguiente diseño de investigación en el que se utiliza la equiparación. Con el fin de animar a los países africanos que parecen encaminarse hacia una mayor democratización, el gobierno de los Estados Unidos decide enviarles asistencia —textos educativos sobre la democracia y materiales similares— a través de un programa denominado «ayuda a la democracia». El investigador quiere estudiar si, en cada país, esa ayuda aumenta el nivel de democracia, si lo disminuye o si no influye en él, pero no puede conceder o suspender la ayuda al mismo país simultáneamente. En consecuencia, opta por un enfoque prospectivo-comparativo: compara los países que están a punto de recibirla con los que no lo están. También decide, con razón, buscar unidades en ambos grupos que sean equiparables en cuanto a los valores de todas las variables de control relevantes, excepto en aquella que le preocupa: el programa de ayuda estadounidense.

El tiempo y los idiomas que conoce condicionan su investigación, de manera que sólo puede estudiar dos países (aunque los problemas que van a mencionarse seguirían apareciendo en un estudio que tuviera un mayor —aunque todavía escaso— número de unidades). Elige un país que recibe bastante ayuda del citado programa norteamericano y otro que apenas la recibe. Acierta al situar su variable dependiente en el incremento del nivel de democracia que se registra en el país desde que comienza el programa de ayuda hasta que se hace el estudio dos años más tarde. Como hay muchas otras variables que podrían correlacionarse tanto con la variable explicativa como con la dependiente, el investigador intenta elegir dos países que sean muy equiparables al respecto, con el fin de evitar el sesgo de la variable omitida.

Podría haber dos variables de control: el nivel educativo del país y el de violencia guerrillera contra el régimen. Ambas pueden producir sesgo si no se controlan, porque las dos se correlacionan tanto con la variable explicativa como con las dependientes (recuérdese el apartado segundo de este capítulo, que versaba del sesgo de la variable omitida). Es probable que los Estados Unidos den más ayuda a aquellos países que tienen un buen sistema educativo (quizá porque éstos podrán establecer mejores relaciones con Washington o porque Estados Unidos impulsa la educación), y, a veces, la enseñanza funciona como una fuerza democratizadora. Del mismo modo, los Estados Unidos prefieren ayudar a los países donde la actividad guerrillera es escasa, y, por supuesto, una amenaza de este tipo hace que disminuya la probabilidad de democratización. Al equiparar en función de estas variables, el investigador espera controlar sus efectos engañosos.

Sin embargo, siempre hay otras variables omitidas que podrían causar sesgo porque se correlacionan tanto con la variable explicativa clave como con la dependiente (y porque, desde el punto de vista causal, son anteriores a aquélla). El problema está en que intentar equiparar dos unidades, si se hace de forma incorrecta o incompleta, puede aumentar la probabilidad de que haya otra variable omitida significativa que se correlacione tanto con la variable explicativa como con la dependiente.

¿Por qué es así? Hay que señalar que, para equiparar países, el investigador tiene que encontrar uno que reciba bastante ayuda y otro que reciba poca. Supongamos que elige dos países que son similares respecto a las otras dos variables, es decir, que tienen un alto nivel educativo y pocas amenazas internas. El resultado sería el siguiente:

País A: mucha ayuda, mucha educación, pacífico.

País B: poca ayuda, mucha educación, pacífico.

Es posible que el país B tenga algo «especial». ¿Por qué no recibe ayuda si tiene condiciones tan favorables? Quizá lo «especial» sea una variable omitida que causará sesgo porque se correlaciona con la variable explicativa y con la dependiente. Un ejemplo de ello podría ser la existencia en B pero no en A de un poderoso ejército que impulse la educación y acabe con los movimientos guerrilleros. Como la fuerza de los militares se correlaciona con la variable dependiente y con la explicativa principal, omitir este factor producirá sesgo. Se puede apreciar que habría existido el mismo problema si la equiparación hubiera partido del otro extremo de las gradaciones de educación y de paz interna. En ese caso, la anomalía sería el país que registra un nivel de educación bajo y mucha violencia y que recibe mucha ayuda. El problema se podría mitigar si la equiparación se realizara utilizando los valores medios de las variables que se basan en la educación y la paz interna. Sin embargo, incluso en este caso, el investigador dispondría de dos países anómalos, cada uno en una dirección opuesta. Lo importante aquí es que, a veces, la equiparación nos induce a buscar observaciones que, en cierto sentido, se apartan de lo que cabría esperar si se dan unos determinados valores en las variables de control y que esa desviación puede deberse a variables omitidas especialmente significativas.

Nótese de qué manera esto tendría que funcionar en el ejemplo de la cárcel. Podríamos intentar equiparar a los prisioneros según ciertas observaciones: que fueran parecidos en cuanto a la procedencia social y económica, la historia de su familia, el expediente escolar y otros factores similares, pero no en el hecho de estar en la cárcel. La mejor equiparación sería la que se basara en personas sin pasado carcelario pero con muchas probabilidades de haberlo tenido, por proceder de un barrio muy pobre, haber dejado el colegio, estar en contacto con el mundo de las drogas, proceder de una familia rota, etc. Cuanto mejor sea la equiparación, más podremos fiarnos

de la relación entre encarcelamiento e ideas políticas. Sin embargo, aquí está de nuevo el problema. Si todo estaba en contra de los no encarcelados, puede que éstos tengan algo especial que les ha mantenido fuera de la prisión —quizá una arraigada práctica religiosa— y que se correlaciona tanto con la variable explicativa (encarcelamiento) como con la dependiente (ideología política).

Hay otra forma de abordar este peligro de la equiparación. Recordemos los dos puntos de vista respecto a la variabilidad aleatoria que se describieron en el apartado sexto del capítulo 2. En la equiparación, tal como la hemos trazado hasta ahora, es posible que planteemos problemas una variable omitida identificable. Sin embargo, aún podríamos sospechar que las dos variables equiparadas en una larga lista de variables de control son «especiales» en algún sentido que no podemos determinar, es decir, que existe una variable omitida desconocida. En esta situación, lo único que podemos hacer es ocuparnos de cómo la aleatoriedad inherente a nuestra variable dependiente influirá en esta observación. Cuanto más se vaya alejando nuestra medida de su verdadero valor, a causa de la variación aleatoria, con más ahínco buscaremos observaciones «inusuales» con el fin de lograr una mayor equiparación entre los grupos y, de este modo, nos arriesgaremos a caer en un sesgo de la variable omitida.

Estas matizaciones *no* tendrían que llevarnos a evitar diseños de investigación que utilicen la equiparación. De hecho, este procedimiento es una de las estrategias más valiosas cuando *n* es pequeño. Simplemente, habrá que ser conscientes de que con la equiparación, al igual que con todas las estrategias que se utilizan cuando *n* tiene un tamaño reducido, se corren riesgos que se hubieran evitado con la aleatorización y con una muestra más grande. Un procedimiento muy fructífero es utilizar la equiparación para elegir estudios de caso pero seleccionar con otros criterios las observaciones de cada uno de ellos.

Si lo que se quiere es evitar el sesgo de la variable omitida, la equiparación tiene que ver con el debate que se mantiene actualmente en la bibliografía sobre política comparada acerca de la conveniencia de seleccionar observaciones lo más parecidas posible (Lijphart, 1971) o lo más diferentes que se pueda (Przeworski y Teune, 1970). Aquí recomendamos un enfoque diferente. Este debate sobre el diseño de investigación, que enfrenta a los partidarios de lo «más parecido» con los de lo «más diferente», se ocupa poco, o nada, de precisar «a qué hay que parecerse». Las etiquetas suelen ser confusas y la polémica no resulta concluyente en estos términos: ninguno de los dos enfoques es siempre preferible. Para nosotros, lo principal a la hora de recoger datos es identificar observaciones que puedan maximizar el control sobre las hipótesis causales. A veces nuestra estrategia conducirá a un plan de investigación que podría calificarse de «diseño de sistemas lo más parecidos posible», otras a lo que se podría denominar «diseño de sistemas lo más diferentes posible». Sin embargo, a diferencia del debate que

contraponen ambas posibilidades, nuestro procedimiento siempre generará datos relevantes para responder a las preguntas del investigador.

En la equiparación, los posibles efectos de las variables omitidas se controlan mediante la selección de observaciones que tienen el mismo valor en tales variables. Por ejemplo, el deseo de mantener constante el mayor número posible de variables contextuales subyace en las opciones de Seymour Martin Lipset (1963, p. 248) al comparar el desarrollo político de los Estados Unidos con el de otras ex colonias británicas de habla inglesa. Lipset indica que los Estados Unidos, Canadá y Australia «... son ex colonias de Gran Bretaña que fijaron fronteras relativamente abiertas y que hoy son estados federales que abarcan continentes». Además, señala muchos otros rasgos comunes que se mantienen constantes: el nivel de desarrollo, el régimen democrático, la similitud de valores, etc.

El estudio que realizó David Laitin (1986) sobre los efectos de las creencias religiosas en la política utiliza una técnica de equiparación especialmente cuidadosa. Eligió un país, Nigeria, en el que eran muy acentuadas tanto la tradición musulmana como la cristiana porque quería comparar las consecuencias que tenían ambas en la política. Sin embargo, las zonas musulmana y cristiana de Nigeria difieren en muchos otros rasgos, además de en el religioso, y hacer caso omiso de ellos habría sido arriesgarse a incurrir en un sesgo de la variable omitida. «En Nigeria, los centros islámicos dominantes se encuentran en los estados del norte, que han tenido durante siglos contacto directo con el mundo musulmán, una historia como países islámicos que precede al dominio británico y una memoria del renacer de la guerra santa a principios del siglo XIX que unificó una zona muy amplia bajo una doctrina islámica ortodoxa. [Por el contrario], hasta finales del siglo XIX las comunidades cristianas no arraigaron [...]. Las escuelas de las misiones trajeron la educación occidental y los empresarios capitalistas animaron a la población a introducir cultivos comerciales y a entrar cada vez más en la economía capitalista mundial» (Laitin, 1986, p. 187).

Laitin se preguntó cómo «se podían controlar las diferencias nacionales, las económicas, las relativas al número de generaciones que se habían visto influidas por una cultura abierta al mundo, las que tenían que ver con las motivaciones para convertirse o las de tipo ecológico; todas ellas diferentes en los baluartes cristianos y musulmanes (1986, pp. 192-193). Optó por elegir una determinada localidad de la región nigeriana de Yoruba, en la que las dos religiones se habían introducido en el mismo grupo nacional más o menos a la vez y donde ambos credos atraían por razones parecidas a los posibles conversos.

La equiparación no es completa ni en el estudio que realizó Kohli sobre los tres estados de la India ni en el análisis de Lipset acerca de las tres ex colonias británicas, ni tampoco en la investigación de Laitin sobre los musulmanes de la tierra yoruba, y nunca podría serlo. Para equiparar es necesario prever y especificar qué posibles variables omitidas relevantes van a

aparecer. Después se controla la investigación seleccionando observaciones cuyos valores no cambian en dichas variables. Es evidente que no podremos saber si hemos cubierto toda la lista de posibles factores de sesgo. Sin embargo, si tenemos ciertos objetivos analíticos —y al evaluar si es adecuado un procedimiento de selección que tenga como fin equiparar hay que relacionarlo con algún objetivo de este tipo—, el control que genere la equiparación aumentará la probabilidad de obtener inferencias válidas.

En suma, el investigador que intente hacer inferencias causales puede seleccionar los casos de dos maneras. La primera es la que utiliza criterios aleatorios para seleccionar y adjudicar valores, que es útil en estudios con un n grande. En tales investigaciones, la aleatoriedad cumple automáticamente la independencia condicional y es un procedimiento mucho más fácil que el de seleccionar observaciones a propósito para cumplir la homogeneidad de las unidades. La aleatoriedad nos garantiza que no se omiten variables relevantes y que la selección de observaciones no se está realizando con criterios que se correlacionan con la variable dependiente (después de controlar las explicativas). Este procedimiento también nos garantiza que los sesgos del investigador no entran en el criterio de selección y que, por tanto, no influyen en los resultados. El segundo método es el de la selección intencionada de observaciones, que recomendamos en estudios que tengan un n reducido. Este tipo de investigaciones, que parten de una selección intencionada para hacer inferencias causales razonables, casi siempre arrojarán inferencias más arriesgadas que las que aplican criterios aleatorios a grandes muestras, y dependerán más de las opiniones previas que tenga el investigador sobre el mundo empírico; en ellas los controles pueden introducir sesgos diversos y sutiles. No obstante, por las razones que se han subrayado, los controles son necesarios en un estudio de n pequeño, y, si son adecuados —si las variables de control se mantienen constantes, quizá mediante la equiparación—, quizá sólo haya que calcular el efecto causal de una única variable explicativa, con lo que aumentaremos el control que tenemos sobre un problema.

7. Observaciones finales

Esperamos que los consejos que hemos dado en este capítulo y en el anterior sean útiles para los investigadores cualitativos, aunque no sean recetas que siempre puedan aplicarse fácilmente. En la realidad, los problemas no suelen venir solos. Por ejemplo, supongamos que un investigador detecta un sesgo de selección menor, un cierto error de medida en la variable dependiente y una importante variable de control que sólo puede medirse ocasionalmente. Si sigue los consejos que se han dado anteriormente para cada caso, tendrá cierta idea sobre lo que hay que hacer. Sin embargo, en este y en otros casos complejos, los que realizan estudios cualitativos tienen que

reflexionar con cuidado sobre los problemas metodológicos específicos que aparecen en sus investigaciones. Puede serles de utilidad remitirse a modelos formales de investigación cualitativa parecidos a los que hemos presentado aquí, pero sólo si están en consonancia con los problemas de su investigación. Gran parte de las ideas que respaldan los modelos formales más sofisticados están en los escritos estadísticos, de manera que no siempre es necesario desarrollar un modelo propio.

El cualitativista, independientemente de que se ayude o no con modelos formales, tiene que atender de manera explícita esas cuestiones metodológicas. Éstas son tan importantes para los investigadores cualitativos que pretenden hacer inferencias causales como para los de orientación cuantitativa.

6. Aumentar el número de observaciones

A lo largo de este libro hemos venido recalcando la importancia crucial que tiene maximizar el control sobre los problemas de la investigación. La principal forma de hacerlo es encontrar tantas consecuencias observables de nuestra teoría como sea posible y hacer observaciones acerca de ellas. Tal como hemos subrayado, lo que puede parecer un estudio de un solo caso, o de unos pocos, quizá contenga en realidad muchas posibles observaciones que sean relevantes para la teoría que se está contrastando y que se encuentren en diferentes niveles de análisis. Al aumentar el número de observaciones, incluso sin recoger más datos, el investigador suele poder transformar un problema insoluble que tiene un diseño de investigación impreciso en algo manejable. En este último capítulo se dan consejos sobre cómo aumentar el número de observaciones relevantes en un estudio social de tipo científico.

Comenzaremos por analizar los problemas inherentes a las investigaciones que se ocupan de una sola observación: el problema que se plantea cuando $n = 1$. Demostraremos que, si realmente hay una sola observación, resulta imposible evitar el problema fundamental de la inferencia causal. Incluso en supuestos ejemplos de comprobación de un único caso, el investigador tiene que examinar, al menos, el pequeño número de observaciones que existen dentro de los «casos» y compararlas. Sin embargo, una inferencia causal puede llegar a apoyarse en una comparación disciplinada que se centre en un reducido número de estudios de caso que ofrezcan observaciones equiparables.

El análisis que hacemos en el apartado 1 de este capítulo de los diseños que se centran en una sola observación puede parecerle pesimista al investigador que utilice el estudio de caso. Sin embargo, como un solo caso puede contener en realidad muchas posibles observaciones, este pesimismo no está justificado, aunque sí está garantizada la búsqueda persistente de más observaciones. Después de criticar los diseños que trabajan con una única observación y, en consecuencia, dar sólidas razones para que se aumente el número de observaciones, pasaremos a analizar cuántas son suficientes para alcanzar niveles de incertidumbre satisfactorios (apartado 2). Para terminar, en el apartado 3 demostraremos que casi todos los diseños de investigación cualitativos pueden reformularse para darles muchas observaciones y que, si el investigador conceptualiza adecuadamente las consecuencias observables que ya tiene, esto suele poder hacerse sin que sea necesario llevar a cabo una costosa recogida de datos adicionales.

1. Inferencias causales con diseños de una sola observación

En una investigación el problema más difícil se plantea cuando el investigador sólo dispone de una unidad para evaluar una teoría causal, o sea, cuando $n = 1$. En este apartado comenzaremos a analizar esta cuestión señalando que es muy improbable que se pueda salir airoso de ella. En primer lugar, estudiamos el argumento de Harry Eckstein en el artículo clásico que dedicó a los estudios de caso cruciales (apartado 1.1). Después, en el apartado 1.2, nos ocupamos de un ejemplo concreto, el razonamiento analógico.

1.1 Estudios de caso «cruciales»

Eckstein ha señalado con contundencia que, cuando no se aclaran las condiciones en las se espera que aparezcan ciertas pautas de comportamiento, resulta imposible comprobar bien o mal esas teorías (Eckstein, 1975). Estamos de acuerdo con él en que los investigadores tienen que esforzarse por encontrar teorías que sean precisas en sus pronósticos y por contrastarlas con datos del mundo real.

Sin embargo, Eckstein va aún más lejos y afirma que si nuestra teoría hace pronósticos precisos se podrá utilizar para fines explicativos un estudio de «caso crucial» que, para él, es una investigación que sólo se basa en una «única medida de cualquier variable pertinente» (lo que nosotros denominamos «observación única»). Lo principal en el capítulo de Eckstein es su idea de que «dos estudios de caso [...] [son] más valiosos cuando [...] se llega al punto en el que se "ponen a prueba" otras posibles teorías» (1975, p. 80). En concreto, señala (1975, p. 127) que «un solo caso crucial puede dejar realmente fuera de combate a una teoría». Para Eckstein, los estudios

de caso cruciales pueden refutar teorías suficientemente precisas con una sola observación. En particular, si un investigador elige un estudio de caso que, en principio, no es probable que concuerde con predicciones teóricas —una observación «muy poco probable»— y resulta que, a pesar de todo, la teoría es correcta, ésta habrá superado una prueba difícil y tendremos razones para suscribirla con una mayor confianza. Por el contrario, si los pronósticos de una teoría que no parece plausible coinciden con las observaciones de otra observación «muy probable», la teoría no habrá superado un examen riguroso pero sí una «prueba de plausibilidad», con lo que quizá merezca la pena seguir analizándola.

El argumento de Eckstein es bastante valioso, especialmente cuando indica a los investigadores que comprendan si han de evaluar sus teorías mediante observaciones «poco probables» o «muy probables». La consistencia de nuestra inferencia, en relación a la validez de la teoría, dependerá en gran medida de lo difícil que haya sido la prueba que ésta haya logrado o no superar. Sin embargo, la defensa que hace Eckstein de una comprobación que se base en una observación crucial no es coherente con el problema fundamental de la inferencia causal. Por lo tanto, si define el «caso» igual que nosotros la observación única, creemos que está equivocado en este punto¹.

Hay tres razones por las que dudamos de que un estudio que se base en una observación crucial pueda servir para los fines explicativos que Eckstein le concede: (1) hay muy pocas explicaciones que sólo dependan de una variable causal; para evaluar el impacto de más de una variable explicativa el investigador necesita haber observado más de una consecuencia; (2) medir la observación crucial es difícil y no del todo fiable; y (3) no es razonable abordar la realidad social partiendo de la base de que es una consecuencia de procesos deterministas, de manera que, aunque las medidas fueran perfectas, habría errores aleatorios.

1. Explicaciones alternativas

Supongamos que comenzamos un estudio de caso con la hipótesis de que un determinado factor explicativo aclara el resultado observado. Sin embargo, durante la investigación descubrimos una posible explicación alternativa. En esta situación, hay que hacer una estimación de dos efectos causales —el de la primera hipótesis y el de la otra explicación— pero sólo disponemos de una observación y, por tanto, está claro que el diseño de la investigación será impreciso (capítulo 4, apartado 1). Además, aunque hiciéramos una equiparación (lo cual suele ser una estrategia valiosa), no podríamos comprobar las explicaciones causales con una única observación. Supongamos que pudiéramos hacer una equiparación perfecta utilizando todas las variables relevantes (circunstancia que es muy improbable en las ciencias sociales). Para poder observar cualquier variación que se produjera en la variable explicativa seguiríamos necesitando comparar, como mínimo, dos unidades. En consecuencia, sería imposible extraer una inferencia causal válida si las hipótesis alternativas sólo se comprobaran con una única comparación.

2. Error de medida

Aunque dispusiéramos de una teoría que hiciera predicciones consistentes y precisas, aún tendríamos que enfrentarnos al hecho de que nuestra medida de esa predicción, como todas las medidas, pudiera contener un error (véase el apartado 1 del capítulo 5). Si sólo tenemos una observación, un error de medida bien podría llevarnos a rechazar una hipótesis verdadera, o a lo contrario. Las teorías precisas pueden necesitar medidas más exactas que las que permite el estado actual de nuestras inferencias descriptivas. Si tenemos muchas observaciones podremos reducir la magnitud e influencia del error de medida mediante la agregación, pero en una única observación siempre existirá la posibilidad de que este tipo de error sea determinante y de que nos lleve a una conclusión falsa.

3. Determinismo

La última razón, y quizá la más decisiva, para considerar inadecuados los estudios que sólo se basan en una consecuencia observable tiene que ver con el grado de determinismo que se da en el mundo. Si éste fuera determinista y la observación produjera una medida que no se correspondiera con la teoría, podríamos afirmar con certeza que dicha teoría es falsa. Sin embargo, en toda teoría social de interés siempre existe la posibilidad de que se omitan variables desconocidas, que podrían conducir a un resultado impredecible aunque el modelo fundamental de dicha teoría fuera correcto. Si sólo existe una consecuencia observada de la teoría causal, careceremos de fundamentos para determinar si la observación confirma o desmiente una teoría o si proviene de algún factor desconocido. Aunque tuviéramos dos observaciones en un experimento perfecto, variaríamos sólo un factor explicativo y produjéramos únicamente una observación para diferenciar dos observaciones que, de otro modo, serían idénticas respecto a la variable dependiente, tendríamos que considerar la posibilidad de que, en nuestro mundo probabilístico, existiera algún factor no sistemático, fruto del azar, que cambiara el efecto causal que se está observando. No importa que el mundo sea inherentemente probabilístico (en el sentido del apartado 6 del capítulo 2) o que, simplemente, no podamos controlar todas las posibles variables omitidas. En cualquier caso, nuestras predicciones sobre las relaciones sociales sólo podrán ser precisas desde un punto de vista probabilístico. En realidad, Eckstein admite que los factores fortuitos influyen en todas las investigaciones:

«La posibilidad de que un resultado se deba al azar no puede descartarse en ningún tipo de estudio; incluso en las grandes investigaciones comparadas es más o menos probable [...]. Por lo tanto, la auténtica diferencia entre los estudios de la observación crucial y los comparados es que en estos últimos, pero no en los primeros, se pueden respetar ciertas convenciones para adjudicar un determinado valor a la posibilidad de que haya resultados fortuitos (por ejemplo, "significativo en el nivel de 0,05")».

Eckstein tiene razón cuando afirma que sólo es habitual admitir que existe una determinada posibilidad de que se produzcan resultados fortuitos en los estudios que uti-

lizan un n grande. Sin embargo, admitirlo tiene la misma importancia en todos los estudios, tanto en los que barajan muchas observaciones como en los que utilizan pocas².

En general, podemos concluir que utilizar una sola observación no es útil para la comprobación de hipótesis o teorías. Sin embargo, es preciso hacer una matización. Aunque tengamos un estudio «puro» de una sola observación, en el que sólo haya realmente una para todas las variables relevantes, ésta podrá ser útil para evaluar explicaciones causales si forma parte de un programa de investigación. Si hay otras observaciones únicas, quizá recogidas por otros investigadores, con las que se pueda comparar, ya no estaremos utilizando una sola observación: esto es precisamente lo que estaremos señalando. No hay que confundir la lógica de la explicación con el proceso de realización del estudio. Si dos investigadores llevan a cabo estudios de una sola observación, puede que —si presuponemos que recogen los datos de manera sistemática y equiparable y que comparten de alguna forma sus resultados— se obtenga una comparación entre dos y una inferencia causal válida. También está claro que los estudios de una sola observación pueden hacer importantes aportaciones al resumen de los pormenores históricos o a la inferencia descriptiva, incluso si no hay comparación (véase el apartado 2 del capítulo 2). Evidentemente, un estudio de caso que, como la mayoría de ellos, contenga muchas consecuencias observables no planteará los problemas que aquí se han señalado.

1.2 Razonamiento analógico

Los peligros de los diseños que no utilizan más que una observación quedan patentes cuando se hace referencia a la forma de equiparación habitual que usan los que elaboran políticas y algunos analistas de esa área cuando intentan comprender acontecimientos políticos, es decir, estamos hablando del razonamiento analógico (véase Khong, 1992). El uso adecuado de la analogía es fundamentalmente el mismo que el de la equiparación que mantiene otras variables constantes. Nuestra hipótesis causal es que si dos unidades son las mismas en todos los aspectos relevantes (esto quiere decir que hemos logrado equipararlas o que, dicho de otro modo, hemos encontrado una buena analogía), unos valores parecidos en las variables explicativas relevantes harán que también se asemejen los de la dependiente. Si nuestra equiparación fuera perfecta y no hubiera error aleatorio en el mundo, sabríamos que la situación de crisis a la que se enfrenta en la actualidad el país B, que es equiparable a la del país A el año pasado, tendría las mismas consecuencias que las que se observaron en este último. Si lo expresamos de este modo se verá que el «razonamiento analógico» puede ser apropiado.

Sin embargo, este tipo de argumentación nunca es mejor que el análisis comparado que conlleva. Al igual que suele ocurrir en los estudios comparados, siempre será mejor (o, en último extremo, no peor) partir de más observaciones para nuestra generalización. Por ejemplo, lo que ocurrió en el país A podría haber sido consecuencia de factores estocásticos que habrían quedado anulados si nuestras predicciones se hubieran basado en las crisis ocurridas en otros cinco países equiparables. Como ocurre en todos los estudios que utilizan equiparaciones, la analogía sólo es tan buena como la propia equiparación. Si ésta es incompleta —si se han omitido variables relevantes—, puede que al calcular los efectos causales cometamos errores. De este modo, al igual que en toda investigación social o pronóstico, es importante aclarar lo más posible el grado de incertidumbre que acompaña nuestra predicción. En general, siempre es preciso ir más allá de una sola observación análoga, al margen de lo similar que nos parezca. Es decir, *el método comparado —en el que se combinan datos de muchas observaciones, aunque algunos de ellos no sean analogías cercanas a la situación que nos ocupa— es, como mínimo, siempre tan bueno como la analogía, y lo normal es que sea mejor*. La razón es sencilla: la analogía utiliza sólo una observación para predecir otra, mientras que el método comparado emplea una combinación bien sopesada de un número mayor de observaciones. Éstas nos ayudarán a realizar un pronóstico más preciso y eficiente, siempre que contengan ciertos rasgos que se parezcan de algún modo —aunque sea poco— al acontecimiento que estamos pronosticando y que utilicemos la información adicional de manera razonable. Por lo tanto, si nos tienta el uso de analogías, tendremos que abordarlas desde un punto de vista comparativo y más general, tal como se analiza más adelante en el apartado 3³.

2. ¿Cuántas observaciones son suficientes?

Al llegar a este punto el investigador cualitativo podría plantearse una cuestión cuantitativa: ¿cuántas observaciones son suficientes? La pregunta tiene importantes consecuencias para la evaluación de los estudios existentes y para el diseño de nuevas investigaciones. La respuesta depende en gran medida del diseño del estudio, de qué inferencia causal se intente calcular y de algunos rasgos del mundo que el investigador no puede controlar.

Aquí respondemos a esta pregunta mediante otro sencillo modelo formal tomado de la investigación cualitativa. Al utilizar el mismo modelo de regresión lineal al que nos remitimos constantemente en los capítulos 4 y 5, centramos nuestra atención en el efecto causal de una variable (x_1). Todas las demás se consideran controles que son importantes para evitar el sesgo de la variable omitida u otros problemas. Se puede expresar fácilmente el número de unidades que se precisan en una determinada situación mediante una fórmula sencilla:

$$n = \frac{\sigma^2}{(1 - R_1^2)S_{x1}^2 V(b_1)} \quad (6.1)$$

cuyo contenido pasamos a explicar.

El símbolo n , por supuesto, representa el número de observaciones sobre las que hay que recoger datos, y en este modelo formal se calcula a partir de σ^2 , $V(b_1)$, R_1^2 y S_{x1}^2 . El significado de cada una de estas cantidades es importante, y todas influyen en el número de observaciones que el investigador cualitativo debe recoger para llegar a una inferencia válida. Hacemos la ecuación (6.1) sin más premisas que las que ya se habían introducido⁴. Ahora las describimos según un orden fijado por la posibilidad de que el investigador pueda influir en ellas: (1) variabilidad fundamental σ^2 , (2) incertidumbre de la inferencia causal $V(b_1)$, (3) colinealidad relativa entre la variable causal y las variables de control R_1^2 y (4) varianza de los valores de la variable causal clave S_{x1}^2 ⁵.

1. Variabilidad fundamental σ^2

Cuanto mayor sea la variabilidad fundamental, o variabilidad no explicada de la variable dependiente (tal como se describió en el apartado 6 del capítulo 2), más observaciones habrá que recoger para hacer una inferencia causal fiable. Esto debería ser relativamente intuitivo, ya que si hay más ruido en el sistema será más difícil encontrar una señal clara mediante un número fijo de observaciones. Recoger datos sobre más unidades puede incrementar nuestro control lo suficiente como para permitimos encontrar pautas causales sistemáticas.

Mediante una analogía directa podemos decir que un estimador más eficiente también necesitará que se recojan más datos. Un ejemplo de esta situación aparece cuando en la variable dependiente hay un error de medida aleatorio (apartado 1.2.1 del capítulo 5). Desde el punto de vista del analista, este tipo de error de medida suele equivaler a una mayor variabilidad adicional, ya que no siempre se puede distinguir entre ambos fenómenos. De este modo, si hay una variabilidad fundamental mayor (o, su equivalente, estimaciones menos eficientes), habrá que recabar más datos.

Aunque el investigador no puede influir en la variabilidad fundamental del mundo, esta información es bastante importante en dos sentidos. En primer lugar, cuanto más sepamos sobre un objeto de estudio, más pequeña será (es de suponer que hasta cierto límite positivo) esta variabilidad fundamental (o no explicada), de manera que no será necesario recoger tantas observaciones para saber algo nuevo. Por ejemplo, si supiéramos mucho acerca de lo que decidió el resultado de ciertas batallas en la guerra de independencia de los Estados Unidos, necesitaríamos un número relativamente pequeño de observaciones (batallas) para hacer una estimación del efecto causal de alguna nueva variable explicativa que hubiera generado una hipótesis.

En segundo lugar, aunque comprender el grado de variabilidad fundamental no nos ayude a reducir el número de observaciones sobre las que tenemos que recoger datos, sí será bastante útil para calibrar con precisión la incertidumbre de las inferencias que se

hagan. Esto tendría que quedar claro con la ecuación (6.1), ya que se puede solucionar la incertidumbre del efecto causal $V(b_1)$ considerándola una función de las otras cuatro cantidades (si conocemos n y el resto de las cantidades, excepto en lo relativo a la incertidumbre de la estimación causal). Esto significa que mediante este modelo formal podemos calcular tanto el grado de incertidumbre de una inferencia causal, utilizando información sobre varias observaciones, como la variabilidad fundamental, la varianza de la variable explicativa causal y la relación que existe entre ésta y las de control.

2. Incertidumbre de la inferencia causal $V(b_1)$

La presencia de $V(b_1)$ en el denominador de la ecuación (6.1) demuestra que es obvio que cuanto más incertidumbre estemos dispuestos a tolerar, menos observaciones habrá que recoger. En aquellas áreas en las que cualquier nueva información sea crucial, quizá podamos hacer aportaciones importantes con un número relativamente pequeño de observaciones. En otros casos en los que ya se haya profundizado más en el tema y un estudio sólo pueda ser relevante con un grado de certidumbre considerable, necesitaremos un número relativamente mayor de observaciones para convencer de la existencia de un nuevo efecto causal (véase el apartado 2.1 del capítulo 1).

3. Colinealidad entre la variable causal y las de control R^2

Si la variable causal no se correlaciona con otras variables que estemos controlando, incluir éstas —lo cual puede ser necesario para evitar el sesgo de la variable omitida u otros problemas— no influirá en el número de observaciones que hay que recoger. Sin embargo, cuanto más alta sea la correlación entre la variable causal y cualquier otra que estemos controlando, más le estará pidiendo a los datos el diseño de investigación y, por lo tanto, mayor será el número de observaciones que haya que recabar para alcanzar el mismo nivel de certidumbre.

Por ejemplo, supongamos que estamos realizando un estudio para comprobar si las mujeres reciben igual salario por el mismo trabajo en algún sector. No tenemos acceso a datos oficiales y, por tanto, sólo podemos hacer entrevistas de tipo informal. Nuestra variable dependiente es el sueldo anual de un empleado, y la explicativa clave, el género. Una de las variables de control importantes es la raza. En un caso extremo, si todos los hombres del estudio fueran negros y todas las mujeres blancas, no tendríamos control sobre la inferencia causal: sería imposible hallar las consecuencias del género después de haber controlado el factor raza. En consecuencia, el género se hace constante en esta muestra; de ahí que éste sea un ejemplo de multicolinealidad, un diseño de investigación impreciso (apartado 1 del capítulo 4); sin embargo, hay que fijarse en lo que ocurre cuando la colinealidad es alta pero no perfecta. Supongamos, por ejemplo, que recogemos información sobre quince empleados y que, de éstos, todos los hombres, excepto uno, son negros y todas las mujeres blancas. En esta situación, el efecto que tiene el género, al controlar el factor raza, se basa completamente en la única observación que no es perfectamente colineal.

Por lo tanto, en la situación general, al igual que en este ejemplo, cuanto más colinealidad haya entre la variable causal explicativa y las de control, más observaciones desperdiciaremos. En consecuencia, serán necesarias más para alcanzar un determinado nivel de incertidumbre. Este asunto tiene una gran importancia práctica para el diseño de investigaciones, ya que suele ser posible seleccionar observaciones para mantener baja la correlación entre la variable causal y las de control. En el ejemplo actual, para reducir esta correlación sólo necesitaríamos entrevistar a un número suficiente de mujeres negras y de hombres blancos.

4. Varianza de los valores de la variable explicativa causal S_{x1}^2

Para terminar, cuanto mayor sea la varianza de los valores de la variable explicativa causal, menos observaciones necesitaremos recabar para alcanzar un determinado nivel de certidumbre en relación a la inferencia causal.

Este resultado, al igual que el último que hemos mencionado, tiene consecuencias prácticas, ya que, si seleccionamos apropiadamente las observaciones, podremos reducir la necesidad de que tengan que ser muchas. Sólo necesitamos centrarnos en elegir aquellas que presenten una amplia gama de valores en relación a la variable causal clave. Si nos interesa saber el efecto que tiene el nivel de estudios medio de una comunidad sobre la delincuencia, lo mejor será elegir algunas comunidades que registren niveles muy altos de este indicador y otras que los tengan muy bajos. Si se hace así, será menor el trabajo necesario para extraer una inferencia causal que tenga un determinado nivel de certidumbre porque habrá que recoger menos observaciones.

El presente modelo formal presupone que el efecto que estamos estudiando es lineal. Es decir, cuanto mayores sean los valores de las variables explicativas, más alto (o más bajo) será el valor esperado de la variable dependiente. Lo mismo ocurrirá si la relación no es lineal pero sí bastante uniforme (o sea, no decreciente). Si, por el contrario, el efecto es claramente no lineal, puede que los valores medios de la variable explicativa presenten un resultado completamente diferente. Por ejemplo, supongamos que un estudio que sólo se basara en los valores extremos de la variable explicativa llegara a la conclusión de que no existe ningún efecto: el nivel educativo de la comunidad no influye en la delincuencia. Sin embargo, de hecho, podría ser que sólo los niveles educativos medios redujeran el grado de delincuencia de una comunidad. En la mayoría de los problemas esta matización no es cierta, pero, al diseñar la investigación, hay que procurar especificar exactamente nuestras premisas.

Si prestamos atención a la variabilidad fundamental, a la incertidumbre, a la colinealidad y a la varianza de los valores de la variable causal, podremos lograr más control con pocas unidades. Sin embargo, no deja de ser razonable hacerse la pregunta que da título a este apartado: ¿cuántas observaciones son suficientes? No podemos responder a esta cuestión con una

respuesta precisa que siempre sea aplicable. Como hemos demostrado con el último modelo formal, la respuesta depende de cuatro informaciones diferentes, que variarán en cada diseño de investigación. Además, la mayoría de las situaciones que son objeto de investigación cualitativa no se ajustarán exactamente a este modelo formal, aunque, en un sentido más general, las intuiciones fundamentales sí sean aplicables.

Cuántas más observaciones, mejor, pero ¿cuántas son necesarias? En la situación más sencilla —en la que hay niveles de variabilidad escasos, varianza alta en la variable causal, ninguna correlación entre ésta y las de control y se requieren niveles bastante bajos de certidumbre— se necesitarán pocas observaciones: probablemente más de cinco pero menos de veinte. De nuevo, la respuesta precisa depende de la exacta especificación del modelo formal y del valor exacto de cada uno de sus componentes. Lamentablemente, la investigación cualitativa, por definición, no suele ser precisa y, en consecuencia, no siempre podemos reducir este aspecto a una sola respuesta.

Por fortuna, sí suele ser posible evitar estos problemas aumentando el número de observaciones. A veces esto supone recoger más datos, pero, como señalamos en el apartado siguiente, suele ser posible reconceptualizar un diseño de investigación cualitativo para sacarle muchas más observaciones y darle así una estructura más consistente. A continuación nos ocupamos de este asunto.

3. Aumentar el número de observaciones a partir de unas pocas

Ya hemos hecho hincapié en las dificultades que conlleva investigar con un pequeño número de observaciones y hemos señalado ciertas medidas que se pueden tomar para mejorar el diseño en esos casos. Sin embargo, quizá el lector se haya dado cuenta de que consideramos todos estos consejos «soluciones de recambio», que son útiles cuando la cantidad de observaciones es limitada pero que no lo son tanto como estrategia para aumentarla⁶. Tal como señalamos, estas soluciones de recambio son valiosas porque a veces no podemos recoger más observaciones del tipo que queremos analizar: quizá sólo haya unos pocos ejemplos del fenómeno que nos interesa o puede que sea demasiado caro o arduo investigar más observaciones que las pocas que ya tenemos. En este apartado analizamos diversas formas de aumentar el número de nuestras observaciones que resultan útiles cuando nos encontramos con lo que parece un número escaso de ellas y no disponemos del tiempo o de los recursos para recoger más. Señalamos varias formas de aumentar el número de observaciones relevantes para nuestra teoría mediante la redefinición de las mismas. Estas estrategias de investigación aumentan n a la vez que mantienen la atención en la búsqueda de pruebas

que constaten o refuten la teoría. Tal como hemos recalcado anteriormente, suelen ser útiles incluso después de que se ha terminado la recogida de datos.

Como señalamos en el apartado cuatro del capítulo 2, para Harry Eckstein (1975) un caso es «un fenómeno del que sólo señalamos e interpretamos una única medida en cada variable pertinente». Como la palabra «caso» se ha utilizado de muchas formas diferentes en las ciencias sociales, preferimos centrarnos en las observaciones, que hemos definido como cada una de las medidas de una variable dependiente en una unidad (y de tantas variables explicativas como se disponga para esa misma unidad). Las observaciones son el componente fundamental de la investigación social empírica: las reunimos con el fin de presentar pruebas que consideramos fiables para evaluar nuestras teorías. En realidad, como señalamos en el capítulo 2, en los proyectos de investigación no se estudian fenómenos globales como Francia, la Revolución francesa, las elecciones estadounidenses de 1992 o la decisión iraquí de invadir Kuwait, sino que lo que se hace es seleccionar aspectos de esos fenómenos —conjuntos de variables explicativas y dependientes— y concretarlos en nuestras teorías. Identificamos unidades en las que funcionan tales variables y hacemos observaciones de éstas en las primeras⁷.

En consecuencia, lo que utilizamos para evaluar nuestras teorías es un conjunto de observaciones referidas a ciertas unidades que, a su vez, tienen que ver con unas variables relevantes. Aquí lo que nos interesa es aumentar el número de observaciones, y todas las formas de hacerlo parten de la teoría o hipótesis que estamos comprobando. Tenemos que preguntarnos cuáles son las posibles consecuencias observables de nuestra teoría o hipótesis y cuántos ejemplos podemos encontrar para contrastarlas. Si necesitamos más observaciones para realizar esta comprobación, hay tres formas de conseguirlas: podemos observar más unidades, medirlas otra vez de diferente manera o hacer ambas cosas (es decir, observar más unidades y utilizar, a la vez, otras medidas). Dicho de otro modo, podemos hacer medidas similares en otras unidades (lo cual describimos en el apartado 3.1), utilizar las mismas unidades cambiando las medidas (apartado 3.2) o cambiar tanto las medidas como las unidades (apartado 3.3). El primer método puede considerarse una reproducción completa de nuestra hipótesis: utilizamos las mismas variables explicativas e independientes con nuevos ejemplos. El segundo conlleva una reproducción parcial de nuestra teoría o hipótesis, en la que se utiliza una nueva variable dependiente con las mismas variables explicativas. Para terminar, en el tercero se plantea una hipótesis nueva (o muy revisada) —que estaba implícita en nuestra teoría original y que utiliza otra variable dependiente— para aplicarla a nuevos ejemplos⁸. Si se utilizan estos métodos se pueden observar muchas consecuencias diferentes de nuestra teoría, incluso en aquellas investigaciones que se denominan convencionalmente «estudios de caso» único. De hecho, en un único caso sue-

de haber múltiples medidas de las variables principales; de ahí que, por definición, contenga múltiples observaciones⁹.

3.1 Las mismas medidas con otras unidades

Obtener más observaciones a partir del mismo método de medida es la forma habitual de incrementar el número de observaciones. Se aplica la misma teoría o hipótesis, más o menos con las mismas variables, a más ejemplos del proceso que describe dicha teoría. Las dos formas principales de encontrar más ejemplos observables del proceso que aborda nuestra teoría se basan en transformaciones «en el espacio» o en el tiempo.

El método fundamental para obtener más observaciones «en el espacio» es buscar otras unidades similares: añadir Pakistán, Bangladesh y Sri Lanka a la India en la propia base de datos. Si se tiene suficiente tiempo, dinero y conocimientos, esta forma de proceder es válida. El trabajo llevado a cabo por Kohli en la India (que se analizó en el apartado 4.3 del capítulo 4) es un ejemplo de esta situación y también pone de manifiesto una de las maneras que él utilizó para superar el problema que planteaba estudiar tres estados indios, que se habían seleccionado en función de valores conocidos de la variable dependiente y de las explicativas. El autor se ocupó de otras dos unidades nacionales. Una de ellas era el Chile de Allende, en el que los programas de ayuda a los pobres fracasaron. Kohli señaló que a este fracaso había contribuido la ausencia de una de las tres características que, para él, llevan al éxito de tales programas (en el caso chileno, la falta de un partido político reformista bien organizado)¹⁰. El otro país que consideró era el Zimbabue de Robert Mugabe, que había llegado al poder, cuando Kohli estaba escribiendo su libro, con un programa de reducción de la pobreza parecido al de Bengala occidental. Los resultados, aunque provisionales, parecían constatar la teoría de Kohli. Estos dos casos los trató por encima, pero no se equivocó al considerarlos consecuencias observables de su teoría.

Sin embargo, no es necesario que rebasemos los límites de la unidad que hemos estado estudiando. Una teoría que, en principio, se centre en un estado-nación puede contrastarse en las unidades geográficas que lo componen, ya sean estados, condados, ciudades, regiones, etc. Evidentemente, esto aumenta el margen de variabilidad tanto de las variables explicativas como de la dependiente. Supongamos que quisiéramos comprobar una teoría de la inestabilidad social que relacionara este problema con los precios agrícolas. Una unidad podría ser la nación denominada «India», pero, si tenemos en cuenta las diferentes partes del país, este caso aportará numerosas observaciones sobre la relación que nos interesa. Sin salir de él podemos aumentar el número de observaciones, buscando reproducciones del proceso que se está estudiando.

Los que estudian las políticas sociales suelen ocuparse de unidades administrativas que, en realidad, son subunidades del estado-nación en el que les quieren comprobar sus hipótesis acerca del origen de diversos tipos de políticas. El análisis que llevó a cabo Kohli de tres estados de la India es un ejemplo de una tendencia habitual en este tipo de estudios, que consiste en comparar estados, ciudades o regiones. Sin embargo, para Kohli, el conjunto de observaciones original eran los tres estados indios. Como hemos señalado, el autor los seleccionó de forma que no se podían utilizar para comprobar su hipótesis sobre el efecto de la estructura del régimen en las políticas contra la pobreza en la India. Sin embargo, igual que utilizó otras naciones como unidades de observación, también solucionó gran parte del problema que había en su primera selección de unidades mediante el uso de subunidades. Partiendo de los tres estados indios con los que había comenzado, descendió a otro nivel de observación para aplicar su hipótesis a los *panchayats* (consejos de gobierno locales de los distritos, bloques y pueblos), que son subunidades de los estados. Los *panchayats* diferían considerablemente en cuanto al compromiso de sus dirigentes con las políticas contra la pobreza y a la estructura organizativa local. De este modo, se pudo comprobar el impacto que tenía esa diversidad en los resultados de la política que el autor utilizaba como variables dependientes.

Las subunidades que aportan más observaciones no tienen por qué ser de tipo geográfico. Las teorías que son aplicables al estado-nación también pueden contrastarse en organizaciones gubernamentales o en el marco de determinadas decisiones, lo cual puede hacerse sin tener que ir a otro país. En un trabajo de Verba y otros autores (1993) se puede apreciar cómo se buscan más consecuencias observables de la propia hipótesis en otras unidades no geográficas. En el ejemplo que presentamos en el apartado 4 del capítulo 5, estos investigadores explicaban que el hecho de que los afroamericanos aprendan más habilidades sociales que los latinos se explica por el tipo de iglesia a la que acuden unos y otros; es probable que el primer grupo vaya a iglesias protestantes que tienen una organización de tipo comunitario, mientras que el segundo suele acudir a congregaciones católicas con un funcionamiento jerárquico. Los autores señalan que, si su hipótesis sobre la influencia que tiene la organización de la iglesia es correcta, el mismo tipo de diferencia que aparece entre fieles católicos y protestantes debería apreciarse al comparar otras unidades religiosas, especialmente las sectas protestantes que difieren en cuanto a su forma de organizarse. Les parece que los episcopalianos, cuya iglesia tiene una estructura jerárquica, son bastante parecidos a los católicos en cuanto a la adquisición de habilidades sociales a través de su congregación. El hecho de que los episcopalianos suelen tener más estudios y recursos que, por ejemplo, los baptistas —aunque en su iglesia demuestran menos sus habilidades sociales— concede un mayor grado de control a la confirmación de esta hipótesis causal.

Hay que tener cuidado al decidir si las nuevas unidades son apropiadas para reproducir nuestra hipótesis, es decir, si en ellas puede tener lugar el proceso que implica dicha hipótesis. La validez que tenga la aplicación de la hipótesis a otras clases de unidades dependerá de la teoría e hipótesis que se estén utilizando, así como del tipo de unidades. Si la variable dependiente es la política de protección social, los estados o provincias serán adecuados si pueden implantar ese tipo de políticas, pero si lo que estamos estudiando es la política arancelaria y todas las decisiones al respecto las toma el gobierno central, quizá las unidades estatales o provinciales no sean las indicadas. Del mismo modo, no tendría sentido estudiar los gobiernos locales de la India o de Pakistán para comprobar una teoría relativa a las condiciones en las que una unidad política decide desarrollar armamento nuclear, ya que este tipo de elección la realiza el gobierno central. Para dar otro ejemplo, podemos señalar que se puede comprobar el impacto que tiene el cambio de los precios agrícolas sobre la inestabilidad social en los estados indios, pero que no es factible utilizar varias organizaciones del gobierno indio para verificar esta relación, ya que este proceso no tiene lugar dentro de ellas. Dicho en pocas palabras, el hecho de que las subunidades sean lugares apropiados para observar una teoría «en acción» depende de esa misma teoría. Ésta es la razón por la que aconsejamos que se comience por elaborar una lista de las consecuencias observables de nuestra teoría y no buscando montones de posibles unidades, independientemente de la teoría que se utilice. Sólo podremos elegir las unidades de estudio una vez que hayamos especificado cuál es la teoría.

Un enfoque alternativo es considerar observaciones a lo largo del tiempo. La India actual y la de hace una década pueden ser dos ejemplos del proceso que nos interesa. De hecho, en la mayoría de las investigaciones que se consideran «estudios de caso» hay múltiples medidas de una hipótesis a lo largo del tiempo.

Creemos que aconsejar que se aumente el número de observaciones buscando más ejemplos en las subunidades o teniendo en cuenta más situaciones a lo largo del tiempo es una de las indicaciones más útiles que podemos dar a los investigadores cualitativos. Soluciona el problema que plantea un n pequeño porque lo aumenta sin que sea necesario, por ejemplo, viajar a otro país o analizar una decisión completamente nueva. Sin embargo, es un consejo que hay que seguir con cuidado: la teoría o hipótesis tiene que ser aplicable a la nueva situación, es decir, la subunidad debe contener realmente una consecuencia observable de dicha teoría. No tiene por qué ser (ni siquiera aproximadamente) la consecuencia observable que nos interese en ese preciso momento; siempre que sea una consecuencia de la misma teoría, los datos que organicemos nos darán un mayor control sobre la inferencia causal.

Hay que ser consciente de otro problema. Queremos utilizar estos nuevos ejemplos para someter nuestra teoría a otras pruebas, pero las subuni-

dades o los diversos ejemplos que hemos encontrado a lo largo del tiempo quizá no sean comprobaciones *independientes* de dicha teoría. De este modo, como reconoce George (1982, pp. 20-23), cada nuevo «caso» no nos aportará tanta información sobre el problema como lo haría si las observaciones fueran independientes. El hecho de que las observaciones dependan unas de otras no descalifica las comprobaciones, a menos que esa dependencia sea perfecta, es decir, a menos que se puedan predecir totalmente los nuevos datos a partir de los existentes. Si no se da este caso improbable, al menos sí habrá nueva información en los nuevos datos que nos ayude a analizarlos. Estas nuevas observaciones, que se basan en información no independiente, no aportarán tanta información como las que son completamente independientes, pero, de todos modos, pueden ser útiles.

Esta conclusión tiene dos consecuencias prácticas. En primer lugar, al barajar observaciones parcialmente dependientes tenemos que tener cuidado de no exagerar la certeza de las conclusiones. Especialmente, no hay que considerar que esos datos nos ofrecen tantas observaciones como tendríamos si las observaciones fueran independientes. En segundo lugar, es preciso analizar con precaución por qué las observaciones dependen unas de otras. Con frecuencia, la dependencia procederá de una o de algunas variables omitidas interesantes que pueden inducir a confusión. Por ejemplo, supongamos que nos interesa la participación política de los ciudadanos en los condados de los Estados Unidos. Quizá los condados contiguos no sean independientes porque la gente va a trabajar cada día de unos a otros, por la movilidad residencial o por los valores socioeconómicos similares que tienen las personas que viven en ellos. Recabar datos en condados vecinos sí aportará más información al estudio, aunque no tanta como se podría esperar de condados que fueran completamente independientes de aquellos en los que ya habíamos recogido datos.

Para dar otro ejemplo, pensemos en la relación que existe entre el cambio de los precios agrícolas y la inestabilidad social. Podríamos comprobar si se da esta relación en varios estados indios, calculando estos precios y la inestabilidad social. Sin embargo, los estados no son unidades experimentales aisladas. Los valores de la variable dependiente pueden verse afectados no sólo por los que adopten las variables explicativas dentro de cada unidad, sino por los de las omitidas que están fuera. En un estado el detonador de la inestabilidad social pueden ser los precios agrícolas (tal como predice la teoría), pero esa alteración del orden social puede influir directamente en que se genere una situación parecida en un estado contiguo (lo cual hará que ésta sólo sea una demostración parcialmente independiente de nuestra teoría). Sería posible enfrentarse al problema controlando adecuadamente esta propagación. Algo similar puede ocurrir con la influencia que ejerce cierto período sobre otro posterior. Podríamos reproducir nuestro análisis en la India una década después, pero la inestabilidad social del primer período quizá tuviera un efecto directo en la del segundo.

Estos ejemplos ponen de manifiesto que la reproducción de un análisis sobre nuevas unidades no siempre produce un nuevo estudio importante. Se pueden utilizar las observaciones que estén dentro del estudio actual y que tengan la misma forma que las ya utilizadas para comprobar la hipótesis. De este modo, el investigador del «estudio de caso» podrá darse cuenta de que existen muchas más observaciones de las que pensaba¹¹.

3.2 Nuevas medidas con las mismas unidades

Se pueden obtener más ejemplos para comprobar una teoría o hipótesis con la misma unidad de observación pero cambiando la variable dependiente. Para utilizar este método hay que buscar muchos efectos de la misma causa, lo cual constituye una técnica muy convincente para comprobar una hipótesis. Una vez más, partimos de una teoría o hipótesis, que suponemos es correcta, y nos preguntamos: aparte de la variable dependiente actual, ¿en qué otras cosas cabe esperar que influyan nuestras variables explicativas? Este ejercicio nos puede sugerir indicadores alternativos de la variable dependiente. En el capítulo 1 señalamos que una determinada teoría acerca de la extinción de los dinosaurios estaba relacionada con la composición química de las rocas. De ahí que, incluso una teoría causal que explique un acontecimiento prehistórico singular, tenga múltiples consecuencias observables que puedan evaluarse.

En el ejemplo de las fluctuaciones de los precios agrícolas y la inestabilidad social puede que hayamos medido el segundo factor utilizando la cantidad de desórdenes públicos. Además de la inestabilidad social, podríamos preguntarnos qué más cabría esperar si la teoría es correcta. Quizá haya otras medidas válidas de ese desorden social (comportamientos desviados de uno u otro tipo). Esta investigación podría llevarnos a la hipótesis de que hay otras variables que se verían afectadas, como son el comportamiento electoral, la inversión empresarial o la emigración. El mismo proceso que hace que la fluctuación de los precios agrícolas genere inestabilidad podría vincular el primer factor con los otros resultados mencionados.

La obra de Putnam (1993) sobre el impacto de los recursos sociales en el comportamiento de los gobiernos locales en Italia participa de un enfoque similar. El comportamiento de las regiones no es una medida única. En realidad, Putnam utiliza una amplia gama de variables dependientes para explicar de dónde surge la eficiencia en el comportamiento democrático de las regiones italianas. Dispone de doce indicadores de comportamiento institucional que, ocupándose de las políticas, pretenden medir su proceso de elaboración, las declaraciones que suscitan y su puesta en práctica. Además, utiliza encuestas para calibrar cómo evalúan los ciudadanos la actuación del gobierno. Cada una de estas medidas representa una consecuencia observable de su teoría.

Cómo hemos indicado anteriormente, la utilización de unidades administrativas subnacionales para estudiar una determinada política arancelaria no sería apropiada si este tipo de impuestos los fijara el gobierno central. Aunque las variables explicativas —por ejemplo, el tipo de sector industrial o de producto agrícola— varíen en diferentes estados o provincias, el proceso por el que se fijan los niveles arancelarios (que es lo que le interesa a la hipótesis que se está comprobando) no tiene lugar dentro de las unidades subnacionales. Sin embargo, si la variable dependiente pasa a ser el voto de los representantes de diversos estados o provincias en cuestiones comerciales o arancelarias, sí será posible estudiar el problema. De este modo, podemos aumentar el número de ejemplos en el que opera el proceso teórico.

3.3 Nuevas medidas con nuevas unidades

También podemos salirnos del conjunto de variables explicativas y dependientes que se han aplicado a un determinado conjunto de variables para buscar otras consecuencias observables que manejen nuevas variables y unidades. Puede que las medidas que utilicemos para comprobar hipótesis, que son fundamentalmente nuevas aunque se deriven de las primeras, sean bastante diferentes de las utilizadas hasta el momento. El proceso que describe la nueva teoría quizá no sea aplicable al tipo de unidad que se estudia, sino a otro (con frecuencia, a una unidad que está en un nivel de agregación inferior o superior). La hipótesis general sobre la relación que existe entre los precios agrícolas y la inestabilidad social puede sugerir otras relativas a la incertidumbre y a la inestabilidad en tipos de unidades diferentes, como son las empresas o los entes públicos. También puede llevarnos a hipótesis que se ocupen del comportamiento de los individuos. En este mismo ejemplo podríamos preguntarnos: «si nuestra teoría acerca del efecto que tienen las fluctuaciones de precios en la estabilidad social (que ya hemos comprobado en diversas unidades políticas) es correcta, ¿qué es lo que esto supone para el comportamiento de las empresas, de las cooperativas agrícolas o de los individuos (quizá en el mismo conjunto de unidades políticas)? ¿Qué consecuencias tendría —si tiene alguna— sobre cómo deciden las organizaciones gubernamentales la asignación de recursos? ¿Qué reacción psicológica cabría esperar que suscitara la incertidumbre en los individuos y qué influencia podría tener en sus posibles comportamientos desviados?».

Este enfoque resulta especialmente útil cuando no hay casos observables de un proceso social potencialmente significativo. En el estudio de la guerra nuclear hay un ejemplo de ello. Como nunca ha tenido lugar un enfrentamiento con este tipo de armamento entre dos potencias que dispongan de él, no podemos observar los efectos de las variables explicativas en el estallido de este tipo de guerra. Supongamos que, según nuestra teoría, la presencia de las armas nucleares en ambos bandos ha evitado una guerra

abierta. Aunque no haya ejemplos observables de nuestra hipótesis fundamental, plantear otra más específica podría conllevar otras posibles observaciones. Por ejemplo, podríamos pensar que nuestra teoría implica que el hecho de que ambos bandos posean armas nucleares tendría que evitar la existencia de *amenazas* serias de guerra abierta. Posteriormente, al estudiar la frecuencia y severidad de las amenazas entre parejas de países con capacidad nuclear y entre los que carecen de ella, y al analizar esas amenazas a medida que la probabilidad de guerra pareciera aumentar durante las crisis, podríamos encontrar, y comprobar, otras consecuencias observables de nuestra teoría.

Para llegar a la nueva teoría o hipótesis, que es diferente de la original pero que parte de ella, suele ser necesario desplazarse a un nivel de agregación más bajo y a un nuevo tipo de unidad: esto no supone trasladarse desde una unidad política como la nación a otro nivel de agregación inferior como es la provincia, sino desplazarse desde esas naciones o provincias hasta los individuos que viven en las unidades o hasta las decisiones que toman dentro de ellas. Cada teoría puede implicar diversas conexiones entre variables que conducen a un determinado resultado; es decir, diversas formas de producirse un fenómeno (Dessler, 1991, p. 345). Antes de diseñar comprobaciones empíricas, quizá tengamos que describir un «mecanismo causal» en el que haya series de hipótesis causales relacionadas que indiquen de qué manera se conectan las variables. Definir esos mecanismos causales diversos y buscarlos puede hacer que encontremos una gran variedad de nuevas consecuencias observables de una teoría (en el apartado 2.1 del capítulo 3 señalamos las diferencias que hay entre el concepto de mecanismo causal y nuestra definición fundamental de causalidad, que es más general).

Este desplazamiento hacia un nuevo tipo de «observación» —otra clase de unidad social, un individuo, una decisión— puede conllevar que se introduzcan variables explicativas no aplicables a la unidad original. Con frecuencia, plantear una hipótesis o teoría sobre unidades políticas supone hacer otra elaboración teórica acerca de cómo se produce el resultado concreto que se observa en el nivel de la unidad. En concreto, la hipótesis que se refiere a ésta puede entrañar otras hipótesis referidas a las actitudes y comportamientos que se registran en el nivel de los individuos que viven dentro de las unidades. Estas segundas hipótesis pueden comprobarse utilizando datos sobre esos individuos. Si nos desplazamos al nivel individual, podremos centrarnos en variables psicológicas o en aspectos de la experiencia o la posición social que carecen de sentido si se aplican a las unidades políticas.

Pensemos en el ejemplo de la relación entre precios agrícolas e inestabilidad social. Podríamos plantear una hipótesis referida a una unidad administrativa, como es una nación o provincia. El ejemplo podría ser el siguiente: cuanto mayor es la fluctuación de los precios agrícolas en una

unidad, más probabilidades hay de que se registre inestabilidad social. A su vez, esta hipótesis sugiere otras acerca de los individuos que viven dentro de esas unidades. Por ejemplo, podríamos plantear que hay una mayor probabilidad de que participen de comportamientos sociales perturbadores los más vulnerables a la fluctuación de los precios (los que cultivan ciertos productos o aquellos que sólo pueden abastecerse adecuadamente si los alimentos son baratos). Para comprobar esta hipótesis quizá tuviéramos que medir factores psicológicos como la alienación o el comportamiento individual desviado.

Los estudios que dan explicaciones culturales a los fenómenos políticos dependen a menudo de este tipo de análisis que se hace en el nivel individual¹². El estudio que realizó Weiner sobre las políticas de educación y de trabajo infantil en la India se basa en una explicación cultural: que la razón de que este país sea casi el único del mundo que no tenga una legislación eficiente para imponer la educación universal y que también carezca de leyes que prohíban el trabajo infantil tiene que ver con valores sociales que comparte el ciudadano medio con las élites gobernantes (Weiner, 1991). La India es un único país, y puede decirse que en el estudio de Weiner n es sólo 1. El autor evita este problema de diversas maneras. Para empezar, compara la India con otros países que han desarrollado un sistema de educación universal. También realiza comparaciones menores entre los estados indios; dicho de otro modo, cambia las unidades. Sin embargo, la hipótesis sobre la cultura y las políticas de la India implica, a su vez, otras hipótesis relativas a los valores de los individuos y a lo que piensan de las políticas. En este sentido, los más importantes son los miembros de las élites que participan en la elaboración de las políticas de educación y de trabajo infantil. De este modo, la principal comprobación a que se somete la hipótesis de Weiner tiene lugar en el plano individual. El autor entrevistó en profundidad a las élites para extraer información sobre lo que pensaban acerca de los valores relativos a la educación y el trabajo infantil. Estas opiniones, además de expresar lo que pensaban los entrevistados de las políticas, constituían las consecuencias observables de la hipótesis global de Weiner.

Esta forma de hacerse con más consecuencias observables de una teoría a partir de unidades que están en un nivel de agregación inferior también puede aplicarse a los análisis de las decisiones. George y McKeown mencionan un enfoque denominado «rastreo de procesos», en el que el investigador se fija con atención en «el proceso decisorio por el que varias condiciones iniciales se transforman en resultados» (George y McKeown, 1985, p. 35)¹³. En lugar de considerar que el resultado final (por ejemplo, de una crisis internacional) es la variable independiente, se construyen otras nuevas. De manera que, por ejemplo, se convierte en una nueva variable cada decisión dentro de una secuencia o cada conjunto mensurable de percepciones que los encargados de decidir tengan de acciones o intenciones ajenas. Este enfoque llega con frecuencia al nivel del actor individual. Una teoría

que vincule las condiciones iniciales con los resultados a menudo implicará que en esos actores se dé un determinado conjunto de motivaciones o percepciones. Por consiguiente, el rastreo de procesos conllevará una búsqueda de pruebas —coherentes con la teoría causal general— relativas a los procesos decisivos que han producido los resultados. Para utilizar este procedimiento quizá se necesite entrevistar a los actores o leer lo que han escrito para explicar las razones de sus actos.

Por ejemplo, en las relaciones internacionales, la cooperación entre estados puede producirse por muchas razones: porque se espere obtener beneficios de la reciprocidad, porque funcione una disuasión con amenazas de destrucción o porque haya unos intereses comunes en cierto conjunto de resultados. En cada uno de estos mecanismos causales participarán muchas variables explicativas, pero en cada uno de los mecanismos posibles el conjunto de variables será diferente, y también lo serán las relaciones entre ellas. Un estudio atento del proceso que lleva a las naciones a cooperar quizá nos permita determinar qué mecanismo causal es más posible que esté funcionando. Esto puede conllevar el estudio, entre otros asuntos, de las motivaciones expresadas por los actores o del tipo de comunicación que hay entre ellos.

Desde nuestra perspectiva, el rastreo de procesos y otras formas de elaborar mecanismos causales aumentan el número de observaciones relevantes desde el punto de vista teórico¹⁴. Estas estrategias vinculan el trabajo teórico y el empírico porque utilizan consecuencias observables de una teoría para indicar nuevas observaciones que habría que hacer para evaluarla. Este método, al aportar más observaciones relevantes para las consecuencias de una teoría, puede ayudarnos a superar los dilemas que plantean los estudios de n pequeño y hacer posible que tanto los investigadores como los lectores tengan más confianza en los hallazgos de las ciencias sociales. En cada sucesión de acontecimientos, el rastreo de procesos produce muchas observaciones. Así ocurre también, dentro de cada unidad política, cuando se analizan actitudes o comportamientos individuales. Además, el investigador controla las variables que son aplicables a todas las observaciones porque se relacionan con la sucesión de acontecimientos o con el conjunto de la unidad. En general, un enfoque que sólo preste atención al resultado final hará que el investigador no disponga más que de un número muy escaso de observaciones para resolver el dilema que se presenta cuando se encuentra o bien sesgo de la variable omitida o imprecisión. Al revisar múltiples observaciones relativas a las actitudes o comportamientos individuales, el investigador puede calibrar qué mecanismos causales se han activado.

No es probable que este tipo de análisis produzca inferencias causales consistentes porque se puede poner en marcha más de un mecanismo y, dentro de cada uno de ellos, quizá no esté clara la fuerza relativa de las variables explicativas. Sin embargo, sí que servirá, en cierta medida, para contrastar hipótesis, ya que una elaboración teórica que explique resultados

también puede tener consecuencias para el proceso que lleva a que éstos se produzcan. Por lo tanto, buscar mecanismos causales aporta observaciones que podrían refutar la hipótesis. Quizá este enfoque también permita al investigador hacer alguna generalización descriptiva sobre la frecuencia con que se activa cada posible mecanismo causal y puede que tales generalizaciones descriptivas sirvan de base para futuros análisis de otros mecanismos causales relacionados y de las condiciones en las que cada uno de ellos se pone en marcha.

Para nosotros, el rastreo de procesos y la búsqueda de los fundamentos psicológicos que sustentan una hipótesis desarrollada en un nivel de agregación superior son enfoques muy valiosos. Sin embargo, constituyen extensiones de la lógica analítica fundamental que hemos venido utilizando y no formas de evitarla. Los estudios de este tipo, si quieren hacer alguna aportación a la inferencia causal, tendrán que enfrentarse a todos sus problemas: la homogeneidad de las unidades, la endogeneidad y el sesgo. Para alcanzar inferencias causales válidas es preciso que, en el nivel del individuo que toma las decisiones, nos planteemos las preguntas que surgen al diseñar una investigación y que las respondamos. Habrá que calibrar con precisión las razones que nos dé ese individuo, seleccionar las observaciones de manera que sean independientes del resultado que se ha producido (si no, tendremos problemas de endogeneidad) y evitar que haya variables omitidas relevantes. También es importante recalcar aquí que los mecanismos causales que se rastreen de este modo tendrán que hacer que nuestra teoría sea más, y no menos, restrictiva: este tipo de técnicas ha de aumentar, y no evadir, las posibilidades de *refutación* de una teoría. En suma, el rastreo de procesos y otros análisis de subunidades son útiles para encontrar hipótesis plausibles sobre mecanismos causales que, a su vez, pueden favorecer la elaboración de generalizaciones descriptivas y allanar el camino hacia la inferencia causal. Sin embargo, este enfoque tiene que enfrentarse a todos los problemas del análisis causal.

4. Observaciones finales

En principio y en la práctica, tanto en los estudios cuantitativos como en los cualitativos se plantean los mismos problemas inferenciales. Las investigaciones que se diseñan para ayudarnos a entender la realidad social sólo pueden lograr su propósito mediante la lógica de la inferencia científica. Esta máxima es aplicable a los estudios cualitativos, cuantitativos, con n grande o pequeño, experimentales, de observación, históricos, etnográficos, a los que utilizan la observación participante y a todos los demás que investigan la sociedad de manera científica. Sin embargo, como este capítulo tendría que haber dejado claro, los problemas fundamentales de las inferencias descriptiva y causal suelen ser más difíciles de evitar cuando se utiliza

un n pequeño que cuando éste es grande. En este libro se han señalado formas de incrementar el número de observaciones para un estudio y también cómo hacer inferencias a partir de una cantidad de observaciones relativamente pequeña.

Tanto los investigadores cuantitativos como los cualitativos pueden hacer más eficiente un estimador aumentando la cantidad de información que se utiliza para resolver un problema —con frecuencia mediante el incremento del número de observaciones (apartado 7.2 del capítulo 2)— y, a veces, recurrir a procedimientos aleatorios de selección y adjudicación de valores con el fin de evitar automáticamente el sesgo. Gran parte de los análisis de este libro pretendían ayudar a los cualitativistas a tener estimadores más precisos, pero las técnicas que hemos señalado son diversas y, con frecuencia, hay que llegar a un equilibrio entre varios objetivos de investigación válidos; de ahí que sea difícil concentrar nuestras indicaciones en enunciados concisos que se correspondan con las ecuaciones formales que prefiere la investigación cuantitativa.

Los investigadores que, en el análisis de los fenómenos sociales, deciden prescindir de los procedimientos cuantitativos formales no pueden hacer caso omiso de los sesgos e ineficiencias que crean los diseños de investigación metodológicamente irreflexivos. Sus temas de estudio son tan importantes, y a veces más, que los que analizan los cuantitativistas. Las inferencias descriptivas y causales que hacen los cualitativistas merecen ser tan sensatas como las de cualquier otro investigador, y para que sean válidas sus autores van a necesitar acercarse más que antes a las cuestiones metodológicas. También deben ser más contenidos al diseñar sus estudios y más explícitos al plantear resultados importantes. Los lectores no tienen por qué reformular los estudios cualitativos publicados para hacerlos científicamente válidos. Si un autor conceptualiza un proyecto de investigación de numerosas consecuencias observables utilizando sólo dos de ellas y doce hipótesis causales, no será responsabilidad de quien lo lea o reseñe el explicar que el diseño implícito del autor era mejor que el explícito. Lo que es aún más fundamental, los investigadores que comprendan y expliquen la lógica de sus análisis producirán estudios más valiosos. Por fortuna, las cuestiones metodológicas que los cualitativistas tienen que comprender son, precisamente, las mismas que el resto de los investigadores científicos debe respetar. Una inferencia válida sólo es posible si se entiende y respeta la lógica interna que subyace en todas las investigaciones sociales de tipo científico.

Capítulo 1

1 Rechazamos el concepto de «semiexperimento» o, al menos, la expresión. Un diseño de investigación o bien conlleva un control por parte del investigador de las observaciones y valores de las principales variables causales (en cuyo caso es un experimento) o no tiene este control (entonces es una investigación de carácter no experimental). Ambos tipos de investigación tienen sus ventajas e inconvenientes, y no puede decirse que uno sea siempre mejor que el otro.

2 Aunque nos ocupemos de casi todas las reglas importantes de la inferencia científica, no puede decirse que éstas sean completas. De hecho, casi todos los filósofos están de acuerdo en que, incluso como principio, es imposible llegar a una lógica inductiva completa y exhaustiva.

3 Sin embargo, una hipótesis alternativa, que postula que la extinción fue producida por erupciones volcánicas, también puede explicar la presencia del iridio, y parece que concuerda mejor con el descubrimiento de que todas las extinciones de especies no ocurrieron simultáneamente.

4 Este dilema no es otro que al que se enfrentan los que estudian las ciencias naturales a la hora de decidir si llevan a cabo investigación aplicada o básica. Por ejemplo, la aplicada que tenga como objeto un determinado fármaco o enfermedad puede mejorar, a corto plazo, la atención médica sin aportar gran cosa al conocimiento general de los mecanismos biológicos subyacentes. La investigación básica puede tener consecuencias opuestas. La mayoría de los investigadores señalarían, al igual que se hace en las ciencias sociales, que esta dicotomía es falsa y que la investigación básica conduce finalmente a eficaces resultados prácticos. Sin embargo, todos estarían de acuerdo en que el mejor diseño de investigación es aquel que de alguna manera se las arregla para participar directamente en la resolución de problemas del mundo real y profundizar en los objetivos de lo escrito en una determinada área científica.

El diseño de la investigación social

5 Éste es el principio de falsabilidad (Popper, 1968); asunto sobre el que hay diversas opiniones en la filosofía de la ciencia. Sin embargo, pocos cuestionan la idea de que las teorías deban enunciarse con la mayor claridad posible, de forma que pueda revelarse que están equivocadas.

6 Probablemente, ésta sea la pregunta que se plantea más a menudo en las entrevistas de trabajo de nuestro departamento y de muchos otros.

7 Esta frase es conocida como el «postulado de simplicidad de Jeffreys-Wrinch». El concepto es similar al de la navaja de Occam.

8 Si hemos elegido un asunto de importancia para el mundo real y/o éste hace alguna aportación a lo escrito en un área de estudio, la naturaleza social de la academia corregirá cualquier posible error: alguien reproducirá nuestro estudio con otro conjunto de datos y demostrará que estábamos equivocados.

9 Nos hemos dado cuenta de que muchos estudiantes de doctorado tienen un miedo innecesario a compartir los datos y la información precisa para reproducir sus resultados. Temen que alguien les robe su preciado trabajo o que incluso demuestre que están equivocados. Son temores habituales, pero pocas veces están justificados. Publicar (o, al menos, enviar copias de lo escrito sobre la investigación a otros académicos) y compartir los datos es la mejor forma de darle credibilidad a nuestra aportación. Además, compartir los datos no hará más que facilitar que otros sigan el camino trazado con la investigación que uno ha comenzado. Una vez que ésta se publique, citarán nuestro esfuerzo e incrementarán nuestra visibilidad y reputación.

10 Podemos comprobar nosotros mismos la fiabilidad midiendo una misma cantidad dos veces y fijándonos en si el resultado es igual. A veces esto parece fácil, como en el caso de hacer exactamente la misma pregunta en diferentes ocasiones durante una entrevista; sin embargo, plantearla una vez puede influir en el entrevistado para que en la segunda ocasión responda de una forma coherente con la anterior, de modo que hay que tener cuidado de que las dos mediciones sean realmente independientes.

11 Ejemplo de ello es la utilización de más de un codificador para la extracción de información sistemática de las transcripciones de entrevistas en profundidad. Si dos personas utilizan las mismas normas de codificación, podemos observar con cuánta frecuencia emiten el mismo juicio. Si no se obtienen medidas fiables, se pueden elaborar normas de codificación más precisas e intentarlo de nuevo. Al final, suele ser posible crear un conjunto de normas para que la aplicación de un mismo procedimiento por diferentes codificadores arroje el mismo resultado.

12 La expresión «falacia ecológica» es confusa porque el razonamiento que lleva de los niveles agregados a los individuales no es ni ecológico ni falaz. No es muy afortunado utilizar el término «ecológico» para describir un análisis de tipo agregado. Aunque Robinson (1990) llegó a la conclusión, en su primer artículo sobre este asunto, de que utilizar el análisis agregado para hacer razonamientos sobre individuos es una falacia, muchos cuantitativistas de las ciencias sociales y también estadísticos reconocen ahora que en niveles agregados de análisis sí existe cierta información sobre los individuos, y se han desarrollado muchos métodos de inferencia «ecológica» no sesgados.

Capítulo 2

1 Tampoco podemos rechazar a Caro por pertenecer a otra profesión: la de periodista/biógrafo, cuyos objetivos son diferentes a los del científico social. Su obra se centra en algunos problemas de los que podría ocuparse un politólogo, como ¿qué lleva al éxito o al fracaso en una campaña electoral?, ¿cuál es el papel del dinero y de la finan-

ciación de la campaña en el éxito electoral?, ¿qué motiva a los que dan dinero para la campaña? El análisis se centra en una determinada candidatura de una circunscripción concreta, pero el problema y las cuestiones que plantea se solapan con lo que es habitual en la ciencia política.

2 Con el fin de completar esta explicación, merece la pena señalar que podríamos imaginar una teoría completamente diferente en la que una contracción del párpado no fuera un guiño pero tuviera, a pesar de todo, un efecto causal en otros actores. Por ejemplo, el tic podría haberse malinterpretado. Si también nos interesara si la persona que contrae el párpado *pretendía* hacer un guiño, tendríamos que buscar otras consecuencias observables de esta misma teoría.

3 La bibliografía sobre los estudios de caso comparados es enorme. Para complementarla, algunas de las mejores obras son las de Eckstein (1975), Lijphart (1971) y Collier (1991).

4 Coombs (1964), por ejemplo, demostró que casi todas las colecciones de datos útiles requieren o implican un cierto grado de teoría, o de «miniteoría». Sin embargo, hay muchos datos cuantitativos, e históricos de tipo cualitativo, que se recogen con el propósito explícito de animar a futuros investigadores a que los utilicen para fines no previstos anteriormente. Quince minutos con el *Statistical Abstract of the United States* (Resumen estadístico de los Estados Unidos) convencerá a la mayoría de este hecho. Las labores de recogida de datos también difieren por el grado de rigidez con que los investigadores siguen sus presupuestos previos.

5 Formalmente, para un conjunto de unidades n en el que se mide la variable (y_1, \dots, y_n) , un estadístico h es una función con un valor real que se define como: $h = h(y) = h(y_1, \dots, y_n)$.

6 Este asunto está muy relacionado con el concepto de diseño de investigación impreciso, que analizamos en el apartado 1 del capítulo 4.

7 Véase Popper (1982) para un libro entero dedicado a la defensa del indeterminismo.

8 Evidentemente, lo mismo puede aplicarse a todas las comunidades que podríamos estudiar.

9 Hay que señalar que la aleatoriedad no tiene que ver exactamente con diferentes semanas reales, ya que tanto los acontecimientos imprevistos como las diferencias sistemáticas podrían explicar las divergencias observadas. Por lo tanto, creamos una situación lo más ideal posible en la que podamos imaginar que ponemos en marcha el mundo de nuevo con características sistemáticas constantes y dejando que varíen los factores imprevistos.

10 Por supuesto, y_2 puede tener mucho interés para los habitantes de la circunscripción 5 en ese año y , por tanto, puede valer la pena estudiar los componentes aleatorio y sistemático de ese acontecimiento. No obstante, siempre tendríamos que intentar distinguir lo aleatorio de lo sistemático.

11 Véase King (1991b) para una elaboración de esta distinción.

12 Los economistas tienden a estar más cerca del punto de vista 1, mientras que los estadísticos están más próximos al 2. El primero también es especialmente común en la rama de la ingeniería que se denomina «control de calidad». Los físicos han llegado a debatir esta distinción dentro de la mecánica cuántica. En esta área, los primeros que propusieron el punto de vista 2 fueron los partidarios de la «teoría de la variable oculta». Sin embargo, las obras más recientes tienden a constatar el punto de vista 1: el mundo físico parece intrínsecamente probabilístico. Todos esperamos la resolución de las numerosas contradicciones que quedan en esta importante teoría y en sus consecuen-

cias para la naturaleza del mundo físico. Sin embargo, esta disputa dentro de la física, aunque solía justificar gran parte de la filosofía de las ciencias sociales, probablemente no influirá en la lógica inferencial de éstas o en la práctica de sus investigaciones.

13 En nuestro lenguaje, como analizaremos más adelante en el apartado 5 del capítulo 3, éste es un problema de *endogeneidad*. El prever la cooperación podría conducir a que se convocaran reuniones en la cumbre, de modo que no serían estas cumbres las que explicarían la cooperación, sino que ésta parecería el resultado de tal previsión, ¡lo cual no sería muy sorprendente si los actores son racionales!

14 Hay que señalar que el estimador puede no estar sesgado pero sí ser incoherente. Por ejemplo, Y_i es un estimador no sesgado de μ , pero es incoherente porque no mejora cuando va disminuyendo el número de unidades (n , en realidad, no cambia en absoluto). Un estimador también puede ser coherente pero estar sesgado. Por ejemplo $\bar{Y} - 5/n$ está sesgado, pero es coherente porque $5/n$ se convierte en cero a medida que n se acerca al infinito.

15 En la segunda línea de este cálculo y en los siguientes presuponemos que no existe una correlación espacial entre los distritos.

Capítulo 3

1 A la vista de la preferencia que muestran algunos científicos sociales por la explicación, en vez de por la «simple descripción», no resulta sorprendente que los que estudian acontecimientos complicados pretendan envolver su trabajo con los atavíos de la jerga explicativa. Temen que si no lo hacen así se considerará que su trabajo es inferior. En esencia, una auténtica explicación siempre se basa en inferencias causales. Creemos que los argumentos de la bibliografía sobre «explicación no causal» tienen una terminología confusa y que, en casi todos los casos, se refieren en realidad a la explicación causal o carecen de coherencia interna. Si la incapacidad de explicar que tienen los científicos sociales no se debe a investigaciones débiles o a una falta de imaginación sino a la naturaleza de los difíciles pero significativos problemas que estudian, esos sentimientos de inferioridad no están justificados. Describir bien acontecimientos importantes es mejor que explicar mal cualquier cosa.

2 Nuestro punto de partida en este apartado es el artículo de Holland (1986) sobre causalidad y lo que él denomina «modelo de Rubin». Holland se basa en la obra de numerosos autores. En esta materia, la de Donald Rubin (1974, 1978) fue la que tuvo una relevancia más inmediata, pero Holland también cita, entre otros, a Aristóteles, Locke, Hume, Mill, Suppes, Granger, Fisher y Neyman. Aquí extendemos la definición de efecto causal de Holland utilizando algunas ideas, que expresaron claramente Suppes (1970) y otros autores, en relación a la «causalidad probabilística». Esta ampliación nos ha parecido necesaria porque ninguno de los enfoques que existen puede definir la causalidad en relación a una sola unidad, permitiendo, a la vez, que se dividan los efectos causales en componentes sistemáticos y no sistemáticos.

3 Para más información sobre este ejemplo, véase Gelman y King (1990). De forma más general, I y N pueden representar el grupo de «tratamiento» y el de «control», o cualquiera de los dos tratamientos que se han experimentado, de hecho o en teoría. Evidentemente, si se utiliza este lenguaje, denominar, dentro de una variable explicativa, a un valor «tratamiento» y al otro «control» es una decisión completamente arbitraria.

4 Jon Elster (1983, pp. 34-36) ha señalado que existen muchas situaciones, como en las que hay un tercer factor que explica tanto la variable aparentemente explicativa como la dependiente, en que «el significado de la causalidad no puede expresarse me-

dante enunciados contrafácticos». Según nuestro lenguaje, Elster no hace más que mencionar problemas habituales en las inferencias que, en cierta medida, son siempre inciertas. Sin embargo, estas dificultades de la inferencia no invalidan una definición de causalidad que se base en contrafácticos. A pesar de sus objeciones, Elster reconoce que este tipo de enunciados «tienen un importante papel en el análisis causal» (Elster, 1983, p. 36); de ahí que, para nosotros, su argumento sea más convincente si se considera como un conjunto de valiosas advertencias contra la utilización descuidada de los contrafácticos que como una crítica de su importancia fundamental para la definición de conceptos en el razonamiento causal.

5 Podemos limitar el sentido de la siguiente ecuación a esta circunscripción sustituyendo « \rightarrow » por « \Leftarrow ».

6 Probablemente, la razón sea que Holland es un estadístico que está muy próximo a una versión extrema de la variación aleatoria según el «punto de vista 2» que se ha descrito en el apartado 6 del capítulo 2. Al explicar la «solución estadística» para el problema de la inferencia causal, se acerca mucho a nuestra definición de efecto causal, pero su planteamiento tiene que ver sobre todo con la utilización de diferentes unidades para resolver el problema fundamental, más que con preservar la definición de causalidad en una sola. En concreto, su operador de valor esperado hace una media de las unidades, mientras que nosotros (como se describe posteriormente) hacemos una media de las reproducciones hipotéticas del mismo experimento en una sola unidad (véase Holland, 1986, p. 947).

7 Como indicamos con más detalle en el apartado 2 del capítulo anterior, esta expresión puede llevar a engaño. Una «variable aleatoria» tiene un cierto componente sistemático y, por tanto, no siempre es del todo impredecible. Por desgracia, este lenguaje tiene un significado específico en estadística, y los conceptos que implica son importantes. La razón primordial para utilizar esta terminología es que la aleatoriedad no significa que «todo vale» o que «puede pasar cualquier cosa». En realidad, hace referencia a un proceso probabilístico muy bien especificado que se extrae de entre los muchos que son posibles. Por ejemplo, el proceso aleatorio que determina qué cara de una moneda queda mirando hacia arriba después de arrojarla al aire es muy diferente al que afecta al crecimiento de la burocracia de la Unión Europea o a las inciertas consecuencias políticas de un cambio en el sistema electoral italiano. La clave de nuestra representación es que todos estos procesos «aleatorios» tienen componentes sistemáticos y probabilísticos.

8 Estas mismas distinciones son objeto de debate. Hay autores que señalan que un sistema presidencial puede ser más representativo, y otros, que los regímenes parlamentarios son más decididos.

9 La distinción entre característica sistemática y no sistemática no está en absoluto clara en todas las ocasiones. La repentina enfermedad de un presidente parece un rasgo no sistemático del sistema presidencial. Por otra parte, la vulnerabilidad general de este tipo de ordenamiento ante las fluctuaciones de salud y personalidad de un único individuo es un efecto sistemático que posibilita la aparición de alguna característica no sistemática.

10 Con frecuencia, esta idea se explica afirmando que no hay variable explicativa que sea necesaria o suficiente para que una dependiente tenga un determinado valor. Sin embargo, ésta es una terminología engañosa porque la distinción entre condiciones necesarias y suficientes desaparece en gran medida cuando concedemos a las causas la posibilidad de ser probabilísticas. Como explica Little (1991, p. 27): «Pensemos en la afirmación de que una mala comunicación entre las superpotencias durante una crisis aumenta la posibilidad de guerra. Ésta es una afirmación probabilística, ya que identi-

ca una variable causal (la mala comunicación) y sostiene que con ella aumenta la probabilidad de que se produzca un determinado resultado (una guerra). Sin embargo, no puede convertirse en una afirmación acerca de las condiciones necesarias y suficientes para que haya una guerra, por lo que es irreductiblemente probabilística.

11 El insistir en la necesidad de diferenciar los componentes sistemáticos de los no sistemáticos en las observaciones que son objeto de inferencia causal refleja nuestra idea general de que el mundo, al menos tal como lo conocemos, es más probabilístico que determinístico. De ahí que tampoco estemos de acuerdo con la premisa de Ragin (1987, p. 15), según la cual las «explicaciones que surgen de aplicar el método comparado no se conciben de forma probabilística porque, si es posible, cada uno de los ejemplos de un fenómeno se examina y explica». Aunque se pudiera hacer una especie de censo con la información de cada uno de los ejemplos de un fenómeno, así como con cada permutación y combinación de los valores de las variables explicativas, el mundo, de todas formas, produciría esos datos mediante algún proceso probabilístico (tal como se definió en el apartado 6 del capítulo 2). Esto parece invalidar el enfoque del «álgebra booleana» de Ragin como forma general de diseñar explicaciones teóricas o extraer inferencias. Para aprender de los datos es necesario utilizar la misma lógica de la inferencia científica que analizamos en este libro. Sin embargo, el enfoque de este autor sí puede ser valioso como un tipo de teoría formal (véase el apartado 5.2 del capítulo 3), ya que posibilita que el investigador concrete una teoría y sus consecuencias mucho más fácilmente que si se careciera de él.

12 Con el fin de evitar la utilización de un término constante, presuponemos que todas las variables tienen una media igual a cero. Esto simplifica la presentación, pero no limita en modo alguno nuestras conclusiones.

13 Se nos podría llamar (¡o acusar de!) «justificacionistas» o incluso «justificacionistas probabilísticos» (véase Lakatos, 1970), pero si nos han de poner una etiqueta, preferimos la más coherente y filosófica de «bayesianos» (véanse Leamer, 1978; Zellner, 1971; y Barnett, 1982). De hecho, nuestros objetivos son lo que más nos diferencia de Popper. Si el objetivo es concretamente el suyo, estamos de acuerdo con el procedimiento; si es el nuestro, quizá él estuviera de acuerdo con nosotros. Sin embargo, creemos que nuestros objetivos están más cerca de los que se plantean en las ciencias sociales y también de los que es más probable alcanzar.

14 Como siempre, cuando modificamos una teoría para que concuerde con los datos que hemos recogido, hay que evaluar dicha teoría (o la parte a la que se refieren nuestros datos) en un contexto diferente o en otro conjunto de datos.

15 Otra formulación de las ideas de Popper es la de que «no es posible demostrar una negación». Para este autor, no es posible porque el hecho de que un resultado concuerde con la hipótesis puede proceder, simplemente, de una comprobación equivocada. Quien intente demostrar negaciones siempre se encontrará con este problema. De hecho, las dificultades no sólo serán teóricas, sino también profesionales, ya que las revistas académicas tienden a publicar más resultados positivos que negativos.

Esto ha llevado a lo que se denomina *problema del archivador*, que donde más claramente se manifiesta es en la bibliografía cuantitativa. Supongamos que no existen pautas en el mundo. De manera que cinco de cada cien comprobaciones a las que se someta cualquiera de ellas se apartarán del 95% que constituye el intervalo de confianza y que, por tanto, generarán inferencias incorrectas. Si partimos de la base de que las revistas académicas publican resultados positivos en vez de negativos, sólo publicarán ese 5% que es «significativo», o sea, únicamente conclusiones equivocadas, por lo que

nuestros archivadores estarán llenos de artículos que llegan a conclusiones acertadas! (véase Iyengar y Greenhouse (1988) para una revisión de la bibliografía estadística que se ocupa de este problema). De hecho, los investigadores conocen bien estos incentivos y es probable que también influyan en su trabajo. Aunque el índice de aceptación en las principales publicaciones académicas sea más o menos del 5%, la situación no es tan grave, aunque no deja de ser un importante problema. Desde nuestro punto de vista, el problema del archivador podría solucionarse si todos adoptaran nuestra posición alternativa. *Un resultado negativo es tan útil como uno positivo; los dos pueden aportar la misma cantidad de información sobre el mundo.* Estaremos sobre seguro siempre que presentemos nuestras estimaciones y midamos la incertidumbre.

16 La regla para hacer las mejores preguntas en una entrevista es casi la misma que se utiliza al explicar: ser lo más concreto posible. No hay que preguntar a un estadounidense blanco y conservador: «¿es usted racista?», sino más bien «¿le importaría que su hija se casara con un hombre negro?». No se debe preguntar a la gente si sabe de política; hay que pedirle que nos diga los nombres del secretario de Estado y del presidente de la Cámara de Representantes. En general, y siempre que sea posible, *no hay que pedir al entrevistado que nos haga el trabajo.* Es mejor no pedir la estimación de un efecto causal; tenemos que preguntar por medidas de las variables explicativa y dependiente para hacer la estimación nosotros. No tenemos que pedir motivaciones, sino hechos.

Esta regla no implica que no haya que preguntarle a alguien por qué hizo tal cosa. De hecho, preguntar las motivaciones suele ser muy útil para generar hipótesis. Las motivaciones que uno admite también pueden constituir un valioso conjunto de consecuencias observables. Sin embargo, hay que interpretar las contestaciones como una reacción del entrevistado a las preguntas del investigador, y no necesariamente como una respuesta correcta. Para que las preguntas de este tipo sirvan de algo, hay que diseñar el estudio de manera que una determinada respuesta obtenida (independientemente de las justificaciones, adornos, mentiras o recuerdos selectivos que nos encontremos) sea una consecuencia observable.

Capítulo 4

1 En econometría y estadística a la bibliografía sobre la «identificación» le interesa determinar cuándo es impreciso el diseño de investigación en la investigación cuantitativa y cómo se puede corregir el modelo o recoger diferentes tipos de datos para afrontar el problema. Véanse Hsiao (1983) y King (1989, apartado 8.1).

2 Subrayamos de nuevo que no hay que confundir aleatoriedad con azar. En este contexto, hacer una selección aleatoria significa que cada unidad potencial tiene la misma probabilidad de ser seleccionada para nuestra muestra y que las sucesivas elecciones son independientes; como cuando extraemos nombres de un sombrero y después volvemos a introducirlos en él. Ésta es sólo la forma más simple de aleatoriedad, pero todas ellas dependen de procesos probabilísticos.

3 Algunos ejemplos pueden encontrarse en: Roth (1988), Iyengar y Kinder (1987), Fiorina y Plott (1978), Plott y Levine (1978) y Palfrey (1991).

4 El criterio de selección (1,1,0) omite el extremo inferior de la escala (la unidad baja) y el segundo (0,1,1) prescinde de la unidad que está en el extremo superior (unidad alta). Sólo el tercer caso, en el que no se selecciona «media», no se correlaciona con la variable dependiente.

5 Éste es un buen ejemplo de lo que caracteriza a la ciencia. Cuando introducimos este sesgo para confirmar nuestra conclusión, no nos comportamos como tiene que ha-

cerlo un científico social, sino como suele ser habitual cuando, al discutir de política, defendemos nuestras profundas convicciones seleccionando con frecuencia ejemplos que refuerzan nuestros argumentos. Al investigar, si es posible, hay que intentar tener todas las observaciones. Si se precisa una selección, hay que procurar obtener las observaciones que sean esenciales para decidir cuál es la pregunta que tiene interés, y no sólo las que concuerdan con nuestra posición.

6 En este apartado, no contemplamos la posibilidad de que ciertos proyectos de investigación que se diseñan para no dejar que cambie en absoluto la variable dependiente sean parte de un programa de estudio mayor y que, por tanto, puedan aportar información útil sobre las hipótesis causales. Explicamos esta cuestión en el apartado 4 del capítulo 4.

7 Porter afirma que tiene muchos ejemplos de países que no han alcanzado el éxito competitivo; sin embargo, éstos entran en su análisis mediante una elección selectiva de anécdotas y no se estudian con métodos parecidos a los de sus primeras diez naciones. Cuando los ejemplos que corroboran nuestro argumento se extraen de forma no sistemática de un infinito abanico de posibilidades que incluye elementos que nos favorecen y otros que no, resulta demasiado fácil hacer el ridículo encontrando relaciones que no existen. No vamos a valorar si las hipótesis de Porter son correctas; sólo queremos señalar que la información necesaria para decidir sobre este punto debe recabarse de manera más sistemática.

8 En general, el sesgo de selección tiene lugar cuando ésta se realiza a partir de la variable dependiente, después de tener en cuenta (o controlar) las variables explicativas. Como una de ellas es el método de selección, lo controlamos y no introducimos sesgo.

9 La inferencia también será un poco menos segura si el abanico de valores de las variables explicativas se limita mediante la selección. Véase el apartado 2 del capítulo 6.

10 Dicho de otra forma diferente, si seleccionamos en función de las distribuciones marginales de la variable dependiente y de las explicativas, podremos seguir aprendiendo algo sobre su distribución conjunta al hacer el estudio.

Capítulo 5

1 Una excepción a esta regla se produce cuando los errores sistemáticos positivos anulan los negativos, pero es más correcto definir este caso inusual como un tipo de error de medida no sistemático.

2 Este error, independientemente de que se produzca por nuestra incapacidad de medir el mundo real con precisión o por la aleatoriedad de la naturaleza, siempre nos planteará una pregunta filosófica que puede responderse de varias maneras (apartado 6 del capítulo 2). Cualquiera que sea la posición que aceptemos, la consecuencia será la misma.

3 Imaginamos de nuevo que los cuatro círculos blancos representan o bien un gran número de observaciones que coinciden exactamente con ellos o bien una escasa variación estocástica.

4 Los lectores versados en estadística reconocerán que esto corresponde a la propiedad de homocedasticidad, o varianzas constante.

5 Estos supuestos de error implican que el valor esperado de la variable dependiente observada es igual al esperado para la verdadera variable dependiente:

$$E(Y) = E(Y^* + U) = E(Y^*) + E(U) = E(Y^*) = \beta X$$

6 Seguimos presuponiendo aquí que cada círculo representa o bien datos en los que apenas hay variación estocástica o bien una situación en la que numerosos puntos coinciden en el mismo lugar. Al igual que en el apartado 1 de este mismo capítulo, el propósito de este supuesto es centrarnos en el problema.

7 Es preciso señalar la diferencia que hay entre los dos casos en los que resulta aceptable omitir una variable. En el primero, donde la variable omitida no tiene relación con la dependiente, no existe sesgo y no perdemos ninguna capacidad de predecir los valores futuros de esta última variable. En el segundo caso, en el que la variable omitida no tiene relación con la independiente pero sí con la dependiente, tampoco aparece sesgo en nuestra estimación de la relación entre la variable explicativa incluida en el análisis y la dependiente, pero sí perdemos cierta capacidad de pronosticar con exactitud los valores futuros de la segunda. Por lo tanto, si el estar ya en el poder no tuviera relación con los gastos de la campaña electoral, omitir este factor no sesgaría nuestra estimación de la relación que hay entre ese gasto y los votos. Pero si nuestro objetivo fuera pronosticar, tendríamos que representar todas las variaciones sistemáticas de la variable dependiente, y omitir el hecho de estar ya en el poder nos lo impediría, porque estaríamos dejando de lado una importante variable causal. Sin embargo, aunque a largo plazo nuestro objetivo fuera explicar el voto de la forma más completa y sistemática, podría resultar difícil confiar en varios efectos causales que estuvieran encuadrados dentro de un solo estudio. Por lo tanto, quizá necesitaríamos centrarnos en un único efecto causal (o sólo en unos pocos), cualquiera que fuera nuestro objetivo último.

8 Para decirlo de una forma más precisa, F es la estimación del coeficiente que se produce cuando se hace una regresión entre X_1 y X_2 .

9 Como esta ecuación sólo es exactamente aplicable cuando hay grandes muestras, en realidad estamos analizando la coherencia en vez de la falta de sesgo (apartado 7.1 del capítulo 2). Para ser más exactos, los términos que aparecen entre paréntesis en la ecuación (5.2) se desvanecen al dividirlos por n , a medida que este valor se acerca al infinito.

10 Merece la pena considerar lo que supone ocuparse del efecto causal estimado de los precios del crudo sobre la actitud de la opinión pública ante la escasez de energía, controlando a la vez la cantidad de información televisiva al respecto. Pensemos en dos descripciones que son importantes porque posibilitan análisis y estudios más profundos de los procesos causales. En primer lugar, este efecto estimado sólo representa la influencia que tiene el elemento de los precios del crudo que afecta directamente a la actitud de la opinión pública ante la escasez de energía, y que es diferente del que influye en la audiencia de manera indirecta cuando cambia la información televisiva. En resumen, es el efecto directo que tiene el crudo sobre la opinión pública y no el indirecto. El efecto total puede encontrarse si no se controla en absoluto la cantidad de información televisiva que hay acerca de la escasez de energía. Este efecto también se puede describir como el que tienen los precios de la energía sobre la variable «opinión pública en relación a la escasez de energía si se registran unos niveles fijos de información televisiva al respecto». Como ejemplo de esta posibilidad, imaginemos un experimento en el que controláramos la información que dan las cadenas de televisión sobre la escasez de crudo, manteniendo esta variable constante mientras varían, lógicamente, los precios del petróleo. Como la cantidad de información es constante en este experimento, se controla este factor sin recurrir a ningún otro procedimiento explícito. Aunque no pudiéramos llevar a cabo un experimento, sí podríamos hacer una estimación del efecto condicional que tienen los precios del crudo sobre lo que piensa la opinión pública de la escasez de energía, controlando la información televisiva.

11 Además, podría interesarnos únicamente el efecto directo o indirecto de una variable o, incluso, qué efecto causal tiene alguna otra en una ecuación. Si es así, es perfectamente razonable realizar varios análisis diferentes con los mismos datos, siempre que entendamos las diferencias que hay en la interpretación.

12 Milner se enfrentó al posible sesgo de la variable omitida, pero su razonamiento era incorrecto: «Al considerar diferentes sectores industriales, en diversos países, [el diseño de investigación] permite que cambien [las variables de control omitidas], a la vez que muestra que el argumento fundamental sigue teniendo base» (1988, p. 15). En realidad, la única manera «de mantener constantes las variables de control» es impedir realmente que varíen. Si, mediante otras teorías plausibles, Milner hubiera detectado la importancia de esas variables, habría podido incorporar un conjunto de observaciones que cambiaban en su variable explicativa clave (grado de dependencia económica del país o sector industrial respecto al exterior), pero no en esas variables de control.

13 Los cualitativistas a veces sí manipulan variables explicativas mediante la observación participante. Incluso las entrevistas en profundidad pueden ser un tipo de experimento si plantean de manera sistemática diferentes cuestiones o cambian ciertas condiciones en diversas entrevistas. De hecho, experimentar puede llegar a ser un problema para los que realizan este tipo de entrevistas, ya que quizá se sientan más cómodos al aplicar «tratamientos» experimentales (hacer determinadas preguntas) a ciertas personas que no han sido seleccionadas de forma aleatoria. Los experimentadores tienen sus propios problemas, y son muchos, pero la endogeneidad no se cuenta entre ellos.

14 Para seguir estrictamente los métodos del capítulo 3, habría que llevar a cabo este experimento varias veces y considerar que la media de ellas es el efecto causal medio del tratamiento aplicado. También podría interesarnos la varianza del efecto causal en este individuo.

Capítulo 6

1 Sin embargo, como señalaremos más adelante, Eckstein parece reconocer la debilidad de un argumento que, en realidad, no le lleva a refutar la única observación, sino a buscar muchas observaciones.

2 El estudio de sociología comparada de Bollen, Entwisle y Alderson (1993) demuestra que prácticamente todos los libros y artículos que ellos analizaron concedían algún papel al azar, incluso aquellos que explícitamente utilizaban el método de la diferencia de Mill.

3 Kahneman, Slovic y Tversky (1982) describen una falacia psicológica del razonamiento que tiene lugar cuando los decisores sometidos a incertidumbre eligen analogías que se basan en hechos recientes o disponibles, lo cual hace que sus valoraciones estén sistemáticamente sesgadas. A este fenómeno lo denominan la «heurística de la disponibilidad». Véase también Keane (1988).

4 Las premisas son que $E(Y) = X_1\beta_1 + X\beta$, $V(Y) = \sigma^2$, no existe multicolinealidad e implícitamente todas las expectativas dependen de X .

5 Técnicamente, σ^2 es la varianza de la variable dependiente, que está condicionada por todas las variables explicativas $V(Y|X)$; $V(b_1)$ es el cuadrado del error típico de la estimación del efecto causal de X_1 ; R_1^2 es R^2 calculada a partir de una regresión auxiliar de X_1 y de todas las variables de control, y S_{e1}^2 es la varianza muestral de X_1 .

6 En la bibliografía relativa al método comparado se plantea habitualmente la conveniencia de aumentar el número de observaciones. Lijphart (1971) insiste de forma especial en este punto.

7 Estamos de acuerdo con lo que señala William Baumol (1990, p. 1715) en relación a la historia económica: «Muchos historiadores económicos se tienden a sí mismos una absurda trampa cuando pretenden explicar completamente una determinada evolución histórica. Quien intente describir las «cinco causas principales» del punto crítico al que llegó Gran Bretaña a finales del siglo XIX o las de la depresión económica europea de 1847 estará embarcándose en una empresa imposible. Las ciencias naturales, a pesar de sus muchos avances y conocimiento acumulado, siguen basándose en gran medida en experimentos controlados que, como tales, se centran en la influencia simultánea de una o varias variables. En realidad, los científicos buscan derivaciones parciales, más que explicaciones globales para fenómenos complejos de la realidad».

8 También podemos mantener la misma variable dependiente pero cambiando las explicativas. Sin embargo, en la mayoría de las situaciones, esta estrategia se utiliza para evitar el error de medida, utilizándose medidas múltiples de la misma variable explicativa subyacente.

9 A veces, los investigadores realizan estudios que se describen como reproducciones de otros anteriores y en los que no se incluyen nuevas observaciones. Fundamentalmente, lo que reproducen —o intentan reproducir— son estudios ajenos, con el fin de comprobar si los resultados se repiten. Un cuantitativista intentará repetir el análisis de un estudio anterior utilizando los mismos datos. Un historiador revisará las fuentes utilizadas por otro historiador. Un etnógrafo escuchará las entrevistas grabadas para comprobar si las conclusiones originales eran sensatas. Este tipo de actividad es muy útil porque las pruebas científicas deben ser reproducibles; sin embargo, no es lo que planteamos en este apartado, ya que no conlleva la utilización de nuevas observaciones.

10 Ciertas fuerzas externas también favorecieron el fracaso de Allende, pero Kohli concede más importancia a las internas.

11 Los cuantitativistas han desarrollado una enorme colección de elaboradas técnicas estadísticas para analizar datos en los que aparecen las denominadas propiedades de las series temporales o de la autocorrelación espacial. No sólo consiguen corregir estos problemas, sino que han logrado sacar una extraordinaria información de los datos. Véanse Granger y Newbold (1977), Anselin (1988), Beck (1991) y King (1989, 1991c).

12 La utilización de la «cultura» como variable explicativa en las investigaciones sociales es objeto de un gran debate, pero no es el tema de este libro. Sólo señalaremos que las explicaciones de este tipo deben someterse a las mismas comprobaciones lógicas y de medida que se aplican al resto de las investigaciones.

13 Donald Moon denomina *explicación de los fundamentos* a una versión de este enfoque o, por utilizar una expresión de otros autores, análisis de las razones (Moon, 1975).

14 Lo que George y McKeown denominan «explicación dentro de la observación» constituye, siguiendo a Eckstein, una estrategia de redefinición de la unidad de análisis, con el fin de aumentar el número de observaciones. George y McKeown (1985, p. 36) afirman que en los estudios de caso «el comportamiento del sistema no se resume en un solo punto de datos, sino en una serie de puntos o de curvas que se van trazando a través del tiempo». En nuestra terminología, que tomamos de Eckstein (1975), este método consiste en aumentar el número de observaciones, ya que una sola observación se define como «un fenómeno del que sólo señalamos e interpretamos una única medida en cada variable pertinente».

Bibliografía

- Achen, Christopher H. (1986): *Statistical Analysis of Quasi-Experiments*. Berkeley, University of California Press.
- , y Duncan Snidal (1989): «Rational Deterrence Theory and Comparative Case Studies». *World Politics* 41, núm. 2 (enero), pp. 143-169.
- Álvarez, Walter, y Frank Asaro (1990): «An Extraterrestrial Impact». *Scientific American* (octubre), pp. 78-84.
- Anselin, Luc (1988): *Spatial Econometrics: Methods and Models*. Boston, Kluwer Academic Publishers.
- Barnett, Vic (1982): *Comparative Statistical Inference*. 2.ª ed., Nueva York, Wiley.
- Baumol, William J. (1990): «St. John versus the Hicksians, or a Theorist Malgré Lui?». *The Journal of Economic Literature* 28, núm. 4, pp. 1708-1715.
- Beck, Nathaniel (1991): «Alternative Dynamic Structure». *Political Analysis* 3, pp. 51-87.
- Becker, Howard S. (1966): «Whose Side Are We On?». *Social Problems* 14, pp. 239-247.
- , y Charles C. Ragin (1992): *What Is a Case? Exploring the Foundations of Social Inquiry*. Nueva York, Cambridge University Press.
- Blainey, Geoffrey (1973): *The Causes of War*. Nueva York, Free Press.
- Bollen, Kenneth A., Barbara Entwisle y Arthur S. Alderson (1993): «Macrocomparative Research Methods». En Judith Blake (ed.), *Macrocomparative Research Methods*, Palo Alto, Calif. Annual Reviews, Inc.
- Cain, Bruce, John Ferejohn y Morris Fiorina (1987): *The Personal Vote: Constituency Service and Electoral Independence*. Cambridge, Harvard University Press.
- Caplow, Theodore, Howard M. Bahr, Bruce A. Chadwick y Dwight W. Hoover (1983a): *All Faithful People: Change and Continuity in Middletown's Religion*. Minneapolis, University of Minnesota Press.

- (1983b): *Middletown Families: Fifty Years of Change and Continuity*. Nueva York, Bantam Books.
- Caro, Robert (1983): *The Years of Lyndon Johnson*. Nueva York, Vintage Books.
- Collier, David (1991): «The Comparative Method: Two Decades of Change». En Dankwart A. Rustow y Kenneth Paul (eds.), *Comparative Political Dynamics: Global Research Perspectives*. Nueva York, Harper Collins.
- (1993): «The Comparative Method». En Ada W. Finifter (ed.), *Political Science: The State of the Discipline*. Washington, D.C., American Political Science Association.
- Cook, Karen Schweers, y Margaret Levi (eds.) (1990): *The Limits of Rationality*. Chicago, University of Chicago Press.
- Coombs, Clyde H. (1964): *A Theory of Data*. Nueva York, Wiley.
- Courtillot, Vincent E. (1990): «A Volcanic Eruption». *Scientific American* (octubre), pp. 78-84.
- Dahl, Robert (1961): *Who Governs? Democracy and Power in an American City*. New Haven, Yale University Press.
- Dessler, David (1991): «Beyond Correlations: Toward a Causal Theory of War». *International Studies Quarterly* 3, núm. 35 (septiembre), pp. 337-355.
- Dewald, William G., Jerry G. Thursby y Richard G. Anderson (1986): «The Journal of Money, Credit and Banking Project». *American Economic Review* 76, núm. 4 (septiembre), pp. 587-603.
- Diamond, Larry, y Marc F. Plattner (eds.) (1993): *The Global Resurgence of Democracy*. Baltimore, Johns Hopkins University Press.
- Duneier, Mitchell (1993): *Slim's Table*. Chicago, University of Chicago Press.
- Easton, David (1965): *A Systems Analysis of Political Life*. Nueva York, Wiley.
- Eckstein, Harry (1969): «Authority Relations and Governmental Performance». *Comparative Political Studies* 2, pp. 269-325.
- (1975): «Case Study and Theory in Political Science». En Fred I. Greenstein y Nelson W. Polsby (eds.), *Handbook of Political Science*, vol. 1, *Political Science: Scope and Theory*. Reading, Mass., Addison-Wesley.
- Elster, Jon (1983): *Explaining Technical Change: A Case Study in the Philosophy of Science*. Nueva York, Cambridge University Press [ed. castellana: *El cambio tecnológico: investigaciones sobre la racionalidad y la transformación social*. Barcelona: Gedisa, 1992].
- Fearon, James D. (1991): «Counterfactuals and Hypothesis Testing in Political Science». *World Politics* 43, núm. 2 (enero), pp. 169-195.
- Fenno, Richard F. (1978): *Home Style*. Boston, Little, Brown.
- Ferejohn, John (1993): «Structure and Ideology: Change in Parliament in Early Stuart England». En Judith Goldstein y Robert O. Keohane (eds.), *Ideas and Foreign Policy: Beliefs, Institutions and Political Change*. Ithaca, Cornell University Press.
- Ferguson, Yale H., y Richard W. Mansbach (1988): *The Elusive Quest: Theory and International Politics*. Columbia, University of South Carolina.
- Feynman, Richard P. (1965): *The Character of Physical Law*. Cambridge, Mass., MIT Press [ed. castellana: *El carácter de la ley física*. Barcelona: Orbis, 1987].
- Fiorina, Morris, y Charles R. Platt (1978): «Committee Decisions under Majority Rules». *American Political Science Review* 72, núm. 2 (junio), pp. 575-598.
- Fisher, Ronald A. (1935): *The Design of Experiments*. Nueva York, Hafner Publishing.
- Fogel, Robert William (1989): *Without Consent or Contract: The Rise and Fall of American Slavery*. Nueva York, W. W. Norton.
- Friedrich, Carl J. (1958): «Political Philosophy and the Science of Politics». En Roland Young (ed.), *Approaches to the Study of Politics*. Chicago, University of Chicago Press.
- Fudenberg, Drew, y Jean Tirole (1989): «Noncooperative game theory for industrial organization: an introduction and overview». En Richard Schmalensee y Robert D. Willig (eds.), *Handbook of Industrial Organization*, vol. 1. Amsterdam, North Holland.
- Garfinkel, H. (1964): «Studies of the Routine Grounds of Everyday Activities». *Social Problems* 11, pp. 225-250.
- Geddes, Barbara (1990): «How the Cases You Choose Affect the Answers You Get: Selection Bias in Comparative Politics». *Political Analysis* 2, pp. 131-152.
- Geertz, Clifford (1973): *An Interpretation of Cultures*. Nueva York, Basic Books [ed. castellana: *Interpretación de las culturas*. Barcelona: Gedisa, 1988].
- (1983): «Local Knowledge: Fact and Law in Comparative Perspective». En Clifford Geertz (ed.), *Local Knowledge: Further Essays in Interpretive Anthropology*. Nueva York, Basic Books [ed. castellana: *Conocimiento local: ensayos sobre la interpretación de las culturas*. Barcelona: Paidós, 1994].
- Gelman, Andrew, y Gary King (1990): «Estimating Incumbency Advantage without Bias». *American Journal of Political Science* 34, núm. 4 (noviembre), pp. 1142-1164.
- (1993): «Why Are U.S. Presidential Election Polls So Variable When the Vote Is So Predictable?». *British Journal of Political Science*.
- George, Alexander L. (1982): «Case Studies and Theory Development». Trabajo presentado en el segundo simposio anual sobre el proceso de la información en las organizaciones, Carnegie-Mellon University (octubre), pp. 15-16.
- , y Timothy J. McKeown (1985): «Case Studies and Theories of Organizational Decision Making». *Advances in Information Processing in Organizations* 2, pp. 21-58.
- , y Richard Smoke (1974): *Deterrence in American Foreign Policy*. Nueva York, Columbia University Press.
- Gigerenzer, Gerd, Zeno Swijtink, Theodore Porter, Lorraine Daston, John Beatty y Lorenz Kruger (1989): *The Empire of Chance: How Probability Changed Science and Everyday Life*. Nueva York, Cambridge University Press.
- Gilpin, Robert (1981): *War and Change in World Politics*. Nueva York, Cambridge University Press.
- Goldberger, Arthur (1991): *A Course in Econometrics*. Cambridge, Mass., Harvard University Press.
- Goldstein, Judith, y Robert O. Keohane (eds.) (1993): *Ideas and Foreign Policy: Beliefs, Institutions and Political Change*. Ithaca, Cornell University Press.
- Gould, Stephen J. (1989a): *Wonderful Life: The Burgess Shale and the Nature of History*. Nueva York, Norton [ed. castellana: *La vida maravillosa*. Barcelona: Mondadori, 1995].
- (1989b): «The Horn of Triton». *Natural History* (diciembre), pp. 18-27.
- Granger, G. W. J., y P. Newbold (1977): *Forecasting Economic Time Series*. Nueva York, Academic Press.
- Gulick, Edward V. (1967): *Europe's Classical Balance of Power*. Nueva York, Norton.
- Hall, Peter A. (ed.) (1989): *The Political Power of Economic Ideas: Keynesianism Across Nations*. Princeton, Princeton University Press.
- Halpern, Nina (1993): «Stalinist Political Economy». En Judith Goldstein y Robert O. Keohane (eds.) (1993), *Ideas and Foreign Policy: Beliefs, Institutions and Political Change*. Ithaca, Cornell University Press.

- Hermens, F. A. (1941): *Democracy or Anarchy: A Study of Proportional Representation*. South Bend, Ind., University of Notre Dame Press.
- Hirschman, Albert O. (1970): «The Search for Paradigms as a Hindrance to Understanding». *World Politics* 22, núm. 3 (abril), pp. 329-343.
- Hoffmann, Stanley (1960): *Contemporary Theory in International Relations*. Englewood Cliffs, Prentice-Hall [ed. castellana: *Teorías contemporáneas sobre las relaciones internacionales*. Madrid: Tecnos, 1979].
- Holland, Paul (1986): «Statistics and Causal Inference». *Journal of the American Statistical Association* 81, pp. 945-960.
- Horowitz, Donald (1993): «Comparing Democratic Systems». En Larry Diamond y Marc F. Plattner (eds.), *The Global Resurgence of Democracy*. Baltimore, Johns Hopkins University Press.
- Hsiao, C. (1983): «Identification». En Zvi Griliches y Michael Intriligator (eds.), vol. 1, *Handbook of Econometrics*. Amsterdam, North-Holland.
- Huth, Paul (1988): «Extended Deterrence and the Outbreak of War». *American Political Science Review* 82, núm. 2 (junio), pp. 423-443.
- , y Bruce Russett (1990): «Testing Deterrence Theory: Rigor Makes a Difference». *World Politics* 42, núm. 4 (julio), pp. 466-501.
- Inkeles, Alex, y Peter Rossi (1956): «National Comparisons of Occupational Prestige». *American Journal of Sociology* 61, pp. 329-339.
- Iyengar, Satis, y Joel B. Greenhouse (1988): «Selection Models and the File Drawer Problem». *Statistical Science* 3, núm. 1 (febrero), p. 109-135.
- Iyengar, Shanto, y Donald Kinder (1987): *News That Matters*. Chicago, University of Chicago Press.
- Jeffreys, Harold (1961): *Theory of Probability*. Oxford, Clarendon Press.
- Jervis, Robert (1976): *Perception and Misperception in International Politics*. Princeton, Princeton University Press.
- , Richard Ned Lebow y Janice Gross Stein (1985): *Psychology and Deterrence*. Baltimore, Johns Hopkins University Press.
- Jones, E. L. (1981): *The European Miracle: Environments, Economies, and Geopolitics in the History of Europe and Asia*. Cambridge, Cambridge University Press [ed. castellana: *El milagro europeo*. Madrid: Alianza, 1994].
- Johnston, J. (1984): *Econometric Methods*. 3.ª ed., Nueva York, McGraw Hill [ed. castellana: *Métodos de econometría*. Barcelona: Vicens-Vives, 1994].
- Kahneman, Daniel, Paul Slovic y Amos Tversky (eds.) (1982): *Judgement under Uncertainty: Heuristics and Biases*. Nueva York, Cambridge University Press.
- Katzenstein, Peter J. (1985): *Small States in World Markets: Industrial Policy in Europe*. Ithaca, Cornell University Press [ed. castellana: *Pequeños estados en mercados mundiales: política industrial en Europa*. Madrid: Ministerio de Trabajo y Seguridad Social, 1987].
- Keane, Mark T. (1988): *Analogical Problem Solving*. Chichester, Ellis Horwood.
- Kennedy, Paul (1987): *The Rise and Fall of the Great Powers*. Nueva York, Random House [ed. castellana: *Auge y caída de las grandes potencias*. Barcelona: Plaza y Janés, 1997].
- Keohane, Robert O. (1980): «The Theory of Hegemonic Stability and Changes in International Economic Regimes, 1967-1977». En Ole R. Holsti, Randolph M. Siverson y Alexander L. George (eds.), *Change in the International System*. Boulder, Westview Press.
- (1984): *After Hegemony: Cooperation and Discord in the World Political Economy*. Princeton, Princeton University Press.
- (1988): «International Institutions: Two Approaches». *International Studies Quarterly* 32, p. 379.
- (1989): *International Institutions and State Power: Essays in International Relations Theory*. Boulder, Westview.
- , y Joseph S. Nye, Jr. (1977): *Power and Interdependence: World Politics in Transition*. Boston, Little, Brown.
- Khong, Yuen Foong (1992): *Analogies at War: Korea, Munich, Dien Bien Phu, and the Vietnam Decisions of 1965*. Princeton, Princeton University Press.
- King, Gary (1989): *Unifying Political Methodology: The Likelihood Theory of Statistical Inference*. Nueva York, Cambridge University Press.
- (1993): «The Methodology of Presidency Research». En George Edwards III, John H. Kessel y Bert A. Rockman (eds.), *Researching the Presidency: Vital Questions, New Approaches*. Pittsburgh, University of Pittsburgh Press.
- (1991a): «Constituency Service and Incumbency Advantage». *British Journal of Political Science* 21, núm. 1 (enero), pp. 119-128.
- (1991b): «Stochastic Variation: A Comment on Lewis-Beck and Skalaban's 'The R-Square'». *Political Analysis* 2, pp. 185-200.
- (1991c): «On Political Methodology». *Political Analysis* 2, pp. 1-30.
- Kohli, Atul (1987): *The State and Poverty in India: The Politics of Reform*. Nueva York, Cambridge University Press.
- Kreps, David M. (1990): «Corporate Culture and Economic Theory». En James E. Alt y Kenneth Shepsle (eds.), *Perspectives on Positive Political Economy*. Nueva York, Cambridge University Press.
- Laitin, David D. (1986): *Hegemony and Culture: Politics and Religious Change among the Yoruba*. Chicago, University of Chicago Press.
- Lakatos, Imre (1970): «Falsification and the Methodology of Scientific Research Programs». En I. Lakatos y A. Musgrave (eds.), *Criticism and the Growth of Knowledge*. Cambridge, Cambridge University Press [ed. castellana: *La crítica y desarrollo del conocimiento*. Barcelona: Grijalbo-Mondadori, 1974].
- Lakeman, Enid, y James D. Lambert (1955): *Voting in Democracies*. Londres, Faber and Faber.
- Leamer, Edward E. (1978): *Specification Searches: Ad Hoc Inference With Nonexperimental Data*. Nueva York, Wiley [ed. castellana: *Búsquedas de especificación: inferencia ad hoc con datos no experimentales*. Bilbao: Editorial Española Desclee de Brouwer, 1983].
- (1983): «Let's Take the Con Out of Econometrics». *American Economic Review* 73, núm. 1 (marzo), pp. 31-43.
- Levy, Jack S. (1985): «Theories of General War». *World Politics* 37, núm. 3 (abril), pp. 344-374.
- «Quantitative Studies of Deterrence Success and Failure». En Paul C. Stern, Robert Axelrod, Robert Jervis y Roy Radner (eds.), *Perspectives on Deterrence*. Nueva York, Oxford University Press.
- Liebertson, Stanley (1985): *Making It Count: The Improvement of Social Research and Theory*. Berkeley, University of California Press.
- (1992): «Einstein, Renoir and Greeley: Some Thoughts about Evidence in Sociology». *American Sociological Review* 56 (febrero), pp. 1-15.
- Lijphart, Arend (1971): «Comparative Politics and Comparative Method». *American Political Science Review* 65, núm. 3 (septiembre), pp. 682-698.
- Lindberg, Leon N., y Stuart A. Cheingold (1970): *Europe's Would-Be Polity: Patterns of Change in the European Community*. Englewood Cliffs, Prentice-Hall.

- (1971): *Regional Integration: Theory and Research*. Cambridge, Harvard University Press.
- Linz, Juan J. (1993): «The Perils of Presidentialism». En Larry Diamond y Marc F. Plattner (eds.), *The Global Resurgence of Democracy*. Baltimore, Johns Hopkins University Press, pp. 108-126.
- Lipset, Seymour Martin (1963): *The First New Nation: The United States in Comparative and Historical Perspective*. Nueva York, Basic Books.
- Little, Daniel (1991): *Varieties of Social Explanation: An Introduction to the Philosophy of Social Science*. Boulder, Westview.
- Longino, Helen E. (1990): *Science as Social Knowledge: Values and Objectivity in Scientific Inquiry*. Princeton, Princeton University Press.
- Lowenthal, Abraham F. (1972): *The Dominican Intervention*. Cambridge, Harvard University Press.
- Mankiw, N. Gregory (1990): «A Quick Refresher Course in Macroeconomics». *Journal of Economic Literature* 28, núm. 4 (diciembre), pp. 1645-1660.
- Martin, Lisa L. (1992): *Coercive Cooperation*. Princeton, Princeton University Press.
- Merck & Co., Inc. (1989): *Annual Report*. Rayway, New Jersey, Merck & Co., Inc., 1989.
- Merton, Robert K. [1949] (1968): *Social Theory and Social Structure*. Nueva York, Free Press.
- Mill, John Stuart (1843): *A System of Logic* S. e. [ed. castellana: *Sistema de lógica inductiva y deductiva*. Madrid: Juan Pueyo, 1917].
- Miller, David (1988): «Conjectural Knowledge: Popper's Solution of the Problem of Induction». En Paul Levinson (ed.), *In Pursuit of Truth*. Atlantic Highlands, Humanities Press.
- Milner, Helen V. (1988): *Resisting Protectionism: Global Industries and the Politics of International Trade*. Princeton, Princeton University Press.
- Moe, Terry M. (1990): «The Politics of Structural Choice: Toward a Theory of Public Bureaucracy». En Oliver Williamson (ed.), *Organization Theory: From Chester Barnard to the Present and Beyond*. Nueva York, Oxford University Press.
- Moon, Donald J. (1975): «The Logic of Political Inquiry: A Synthesis of Opposed Perspectives». En Fred I. Greenstein y Nelson W. Polsby (eds.), *Handbook of Political Science*, vol. 1, *Political Science: Scope and Theory*. Reading, Mass., Addison-Wesley.
- Neustadt, Richard E., y Ernest R. May (1986): *Thinking in Time: The Uses of History for Decision-Makers*. Nueva York, Free Press.
- Nye, Joseph S. (1971): *Peace in Parts*. Boston, Little, Brown.
- O'Hear, Anthony (1989): *Introduction to the Philosophy of Science*. Oxford, Clarendon Press.
- Ordeshook, Peter C. (1986): *Game Theory and Political Theory: An Introduction*. Nueva York, Cambridge University Press.
- Palfrey, Thomas R. (ed.) (1991): *Laboratory Research in Political Economy*. Ann Arbor, University of Michigan Press.
- Pearson, Karl (1892): *The Grammar of Science*. Londres, I. M. Dent & Sons, Ltd.
- Plott, Charles R., y Michael E. Levine (1978): «A Model of Agenda Influence on Committee Decisions». *American Economic Review* 68, núm. 1 (marzo), pp. 146-160.
- Popper, Karl R. (1968): *The Logic of Scientific Discovery*. Nueva York, Harper and Row [ed. castellana: *Lógica de la investigación científica*. Madrid: Tecnos, 1985].
- (1982): «The Open Universe: An Argument for Indeterminism». En W. W. Bartley III (ed.), *The Postscript to the Logic of Scientific Discovery*. Totowa, N. J., Rowman and Littlefield.
- Porter, Michael E. (1990): *The Competitive Advantage of Nations*. Nueva York, Free Press [ed. castellana: *Ventaja competitiva de las naciones*. Barcelona: Plaza y Janés, 1991].
- Przeworski, Adam, y Henry Teune (1982): *The Logic of Comparative Social Inquiry*. Malabar, Florida, Krieger Publishing Company.
- Psathas, George (1968): «Ethnomethods and Phenomenology». *Social Research* 35, pp. 500-520.
- Putnam, Robert D., con Robert Leonardi y Raffaella Y. Nanetti (1993): *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton, Princeton University Press.
- Ragin, Charles C. (1987): *The Comparative Method: Moving beyond Qualitative and Quantitative Strategies*. Berkeley, University of California Press.
- Rivers, Douglas, y Morris P. Fiorina (1989): «Constituency Service, Reputation, and the Incumbency Advantage». En Morris P. Fiorina y David Rohde (eds.), *Home Style and Washington Work*. Ann Arbor, University of Michigan Press.
- Robinson, William S. (1990): «Ecological Correlations and the Behavior of Individuals». *American Sociological Review* 15, pp. 351-357.
- Rogowski, Ronald (1987): «Trade and the Variety of Democratic Institutions». *International Organization* 41, núm. 2 (primavera), pp. 203-224.
- Rosenau, Pauline (1990): «Once Again into the Fray: International Relations Confronts the Humanities». *Millennium: Journal of International Studies* 19, núm. 1 (primavera), pp. 83-110.
- Rosenstone, Steven R. (1983): *Forecasting Presidential Elections*. New Haven, Yale University Press.
- Roth, Alvin E. (1988): «Laboratory Experimentation in Economics: A Methodological Overview». *The Economics Journal* 98 (diciembre), pp. 974-1031.
- Rubin, Donald B. (1974): «Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies». *Journal of Educational Psychology* 66, pp. 688-701.
- (1978): «Bayesian Inference for Causal Effects: The Role of Randomization». *The Annals of Statistics* 6, pp. 34-58.
- Russett, Bruce (1978): «The Marginal Utility of Income Transfers to the Third World». *International Organization* 32, núm. 4, pp. 913-928.
- Sanday, Peggy Reves (1983): «The Ethnographic Paradigm(s)». En John Van Maanen (ed.), *Qualitative Methodology*. Sage, Beverly Hills.
- Schumpeter, Joseph A. [1936] (1991): «Can Capitalism Survive?». En Richard Swedberg (ed.), *The Economics of Sociology and Capitalism*. Princeton, Princeton University Press.
- Shepsle, Kenneth A. (1986): «Institutional Equilibrium and Equilibrium Institutions». En Herbert F. Weisberg (ed.), *Political Science: The Science of Politics*. Nueva York, Agathon Press.
- Shively, W. Phillips (1990): *The Craft of Political Research*. 3.^a ed., Englewood Cliffs, Prentice-Hall.
- Simon, Herbert A. (1985): «Human Nature in Politics: The Dialogue of Psychology with Political Science». *American Political Science Review* 79, núm. 2 (junio), pp. 293-305.
- Skocpol, Theda (1979): *States and Social Revolutions*. Cambridge University Press.

- (1984): «Emerging Agendas and Recurrent Strategies in Historical Sociology». En Theda Skocpol (ed.), *Vision and Method in Historical Sociology*. Nueva York, Cambridge University Press [ed. castellana: *Los estados y las revoluciones sociales: un estudio comparativo de Francia, Rusia y China*. México: Fondo de Cultura Económica, 1984].
- Snyder, Glenn H., y Paul Diesing (1977): *Conflict among Nations: Bargaining, Decision Making and System Structure in International Crises*. Princeton, Princeton University Press.
- Snyder, Jack (1991): *Myths of Empire: Domestic Politics and International Ambition*. Ithaca, Cornell University Press.
- Sober, Elliot (1988): *Reconstructing the Past: Parsimony, Evolution and Inference*. Cambridge, MIT Press.
- Suppes, Patrick C. (1970): *A Probabilistic Theory of Causality*. Amsterdam, North-Holland.
- Tawney, R. H. (1935): *Religion and the Rise of Capitalism*. Nueva York, Harcourt, Brace & Co.
- Tilly, Charles (ed.) (1975): *The Formation of National States in Western Europe*. Princeton, Princeton University Press.
- Verba, Sidney (1967): «Some Dilemmas of Political Research». *World Politics* 20 (octubre), pp. 111-128.
- , Kay L. Schlozman, Henry Brady y Norman Nie (1993): «Race, Ethnicity, and Political Resources: Participation in the United States». *British Journal of Political Science* 23, pp. 453-497.
- Waltz, Kenneth N. (1979): *Theory of International Politics*. Reading, Mass., Addison-Wesley.
- Webb, Eugene J., D. T. Campbell, R. D. Schwartz y L. Sechrest (1966): *Unobtrusive Measures*. Chicago, Rand McNally.
- , y Karl E. Weick (1983): «Unobtrusive Measures in Organizational Theory: A Reminder». En John Van Maanen (ed.), *Qualitative Methodology*. Sage, Beverly Hills.
- Weber, Max [1905] (1949): «Critical Studies in the Logic of the Cultural Sciences». En Max Weber (ed.), *The Methodology of the Social Sciences*. Traducido y editado por Edward A. Shils y Henry A. Fluch. Nueva York, Free Press.
- Weiner, Myron (1991): *The Child and the State in India*. Princeton, Princeton University Press.
- Wendt, Alexander (1992): «Anarchy is What States Make of It: The Social Construction of Power Politics». *International Organization* 64, núm. 2 (primavera), pp. 391-426.
- Wolfinger, Raymond, y Steven Rosenstone (1980): *Who Votes*. New Haven, Yale University Press.
- Woods, John, y Douglas Walton (1982): *Argument: The Logic of the Fallacies*. Nueva York, McGraw-Hill Ryerson Ltd.
- Zelditch, Morris Jr. (1971): «Intelligible Comparisons». En Ivan Vallier (ed.), *Comparative Methods in Sociology*. Berkeley y Los Angeles, University of California Press.
- Zellner, Arnold (1971): *An Introduction to Bayesian Inference in Econometrics*. Nueva York, Wiley.
- (1984): *Basic Issues in Econometrics*. Chicago, University of Chicago Press.

Índice analítico

- Achen, Christopher, 140, 145, 255
 acontecimientos únicos, estudio de los, 21-22, 53-55, 104-5
 Alderson, Arthur, 211, 252 n. 2, 255
 Alemania de Weimar, 201-2
 álgebra booleana, enfoque del, 248 n. 11
 alianza, 210
 Álvarez, Luis, 22
 Álvarez, Walter, 22, 255
 análisis bivariable, 184
 análisis contrafáctico, 21, 89-90, 99-100, 203
 análisis de regresión, 109, 140-42, 178-79
 análisis racional, 253 n. 13
 Anderson, Richard, 256
 Anselín, Luc, 253 n. 11, 255
 aportación a los escritos científicos, 27-28
 área de las relaciones internacionales, necesidad de describir en el, 55-56
 armas nucleares, 158-59
 Asaro, Frank, 255
 asignación aleatoria de valores, 105, 126, 207-11, 242
 asignación de valores, 105, 125-26, 207, 242
 aumento de las observaciones, 230-42
 Australia, 217
 autocorrelación espacial, 253 n. 11
 autocorrelación, 253 n. 11
 Bahr, Howard, 255
 Barnett, Vic, 248 n. 13, 255
 Baumol, William, 253 n. 7, 255
 Beatty, John, 257
 Beck, Nathaniel, 253 n. 11, 255
 Becker, Howard, 255
 Blainey, Geoffrey, 255
 Bollen, Kenneth, 211, 252 n. 2, 255
 Cain, Bruce, 197, 255
 Campbell, D. T., 262
 Canadá, 217
 Caplow, Theodore, 37, 255
 carnes rojas en Japón, 43
 Caro, Robert, 47
 casos, 63-64, 127-28, 221-25
 causalidad asimétrica, 100-1
 causalidad múltiple, 98-100

El diseño de la investigación social

- causalidad probabilística, 246 n. 2
 causalidad simétrica, 100-1
 causalidad, 88-101
 Chadwick, Bruce, 255
 Chile, 232
 ciencia como labor social, 19-20
 ciencias naturales, 21
 Cisjordania, 67-70, 106
 codificación de entrevistas, 244 n. 11
 codificación, 168
 coherencia interna, 107, 112, 116-18
 colinealidad, 227-29
Véase también multicolinealidad
 Collier, David, 256
 complejidad de los acontecimientos, 20-21, 54, 104-5
 componente estocástico, 66-74, 91-93, 95, 247 n. 9
 componente no sistemático, 66-74, 91-93, 95, 247 n. 9
 componente sistemático, 45-46, 55-56, 66-74, 91-93, 95, 247 n. 9
Véase también componente estocástico
 comprobación de un único caso, 221-25
Véase también casos; estudios de caso; observaciones
 concisión, 30, 40, 115-16
 concretar, 122
 consecuencias observables, 39, 52, 107, 120-23, 190-91, 236-37
 subunidades de las, 234
 constantes, selección de, 157-59
 control de calidad, 245 n. 12
 control, 134, 178, 197, 207-18
 de lo que se investiga, 211-18
 y problemas del n pequeño, 207-11
 Cook, Karen, 112, 256
 Coombs, Clyde, 245 n. 4, 256
 cooperación entre las superpotencias, 189
 cooperación internacional, 15
 Courtillot, Vincent, 256
 creencias religiosas, 217
 crisis agudas, 145
 cristianismo en Nigeria, 158, 217
 cultura, 120, 239
 política, 203, 239
 cumbre soviética, 72-73
 cursos de contabilidad, 141-42
 Dahl, Robert, 256
 Daston, Lorraine, 257
 datos, recogida de, 62-63
 cuantificación en la, 55
 directrices para mejorarla, 33-36, 57-58
 fiabilidad como objetivo, 36
 garantizar la fiabilidad en la, 36, 162
 nivel de agregación, 60, 128
 para compensar la variabilidad fundamental, 227
 para maximizar el control, 216-17
 proyectos piloto antes de la, 33
 registrar el proceso de, 33-34
 reproducción como objetivo, 36-37
 utilización de modelos formales para la, 62-64, 116-17
 y eficiencia, 227
 y el problema del n pequeño, 226-30
 datos, utilización de los, 37-38
 al resumir, 64-65
 con error de medida, 176-77
 descartar, 165, 195
 para aumentar el control, 41, 115-16, 133, 216-17
 para aumentar el número de observaciones, 230-32
 para comprobar teorías, 24, 30-33, 39-41, 57-60, 108-16
 para garantizar la falta de sesgo, 37-38, 74-76, 106, 110
 para maximizar la eficiencia, 38, 76-85, 106-7, 111
 delincuencia, 171-73, 176-77
 demostración, 210
 descripción detallada, 48-53, 54
 desempleo, consecuencias sobre la delincuencia, 171-73, 176-77
 desfase entre concepto e indicador, 121-22
 Dessler, David, 238, 256
 detalles de los hechos, 48-54
Véase también inferencia descriptiva

- determinista, punto de vista, 70-71, 224-25, 248 n. 11
 Dewald, William, 36, 256
 Diamond, Larry, 256
 Diesing, Paul, 262
 dinosaurios, extinción de los, 21-22, 53, 236
 diseño de investigación impreciso, 128-34, 156
 diseño de la investigación, 24, 28, 128-34, 144, 185, 226-30, 241
 disuasión, teoría de la, 34-35
 Duneier, Mitchell, 48, 256
 Easton, David, 124, 256
 Eckstein, Harry, 48, 63-64, 124, 222-24, 231, 253 n. 14, 256
 educación obligatoria en la India, 76
 efecto causal aleatorio, 91-93
 efecto causal realizado, 90-92, 95
 efecto causal, 92-93, 95-96, 99-100
 modelo formal del, 108-9
 sobreestimarlo, 149
 efecto renta/ingresos, 184-85
 eficiencia, 38, 76-85, 106-7, 111, 161-62, 192-96
 comparada con el sesgo, 80-85
 modelo formal de, 81, 111
 y error de medida, 170, 227
 y recogida de datos, 227, 242
 Einstein, Albert, 17
 elecciones presidenciales, 84
 élites de la política exterior, 135-36
 Elster, Jon, 246-47 n. 4, 256
 empaparse y hurgar, 48-53, 54
 encuestas, 16, 42
 endogeneidad, 105, 118, 197-206, 241, 246 n. 13
 como problema de la variable omitida, 201-2
 como producto natural de los procesos políticos, 210
 modelo formal de, 208
 y elección de observaciones, 202-5
 enfoque comparado, 226
 enfoque globalizador (incluirlo todo), 194
 entrevistas, 249 n. 16, 252 n. 13
 Entwisle, Barbara, 211, 252 n. 2, 255
 equifinalidad, 98
 error aleatorio, 66-74, 91-93, 169-78, 180-81, 95
 error de agrupación, 165
 error de medida, 162-78, 180-81
 modelos formales de, 174-75, 178, 180-81
 no sistemático, 169-78, 180-81
 sistemático, 167-69
 y los estudios de un solo caso, 224
 error medio cuadrático (EMC), 84-85
 errores de predicción, 141
 escarbar en busca de datos, 185
 especificación, 198
 estadísticos, definición de los, 64-65
 estados europeos, 146-47
 Estados Unidos, 217
 estimadores, 194
 estrategias en campañas electorales, 113-15
 estudio de control de caso, 152
 estudios de caso cruciales, 222-25
 estudios de caso, utilización de, 16, 55-57, 63-64, 71-72, 78-79
 cruciales, 221-25
 de la República de Weimar, 201-2
 de un solo caso, 221-25. 253 n. 14
 medidas múltiples a lo largo del tiempo, 234
 y efectos causales constantes, 104
 y homogeneidad de las unidades, 104
 y mecanismos causales, 96-98, 238
 y sesgo de selección, 126
 estudios,
 efecto sobre la renta, 99-100
 en la India, 76
Ética protestante y el espíritu del capitalismo, 198
 Europa del Este, 137-38
 Europa occidental, 146-47
 exclusión de variables relevantes, 72-73, 100, 105, 118, 133-34, 178-93
 modelo formal de, 179-83, 186-87
 y asignación aleatoria de valores, 207
 y endogeneidad, 201-2

y estudios de un solo caso, 223-36
 y selección intencionada, 214-15
 explicación de los fundamentos, 253 n. 13
 explicación dentro de una observación, 253 n. 14
 explicación, 87-88
 falacia ecológica, 40, 244 n. 12
 falsabilidad, 30, 108-10, 112-16, 241
 falta de sesgo, véase sesgo
 Fearon, James, 256
 Fenno, Richard, 49, 256
 Ferejohn, John, 48, 255, 256
 Ferguson, Yale, 256
 fertilizantes, uso de, 158
 Feynman, Richard, 256
 Fiorina, Morris, 197, 249 n. 3, 256
 Fisher, Ronald, 255, 256, 261
 Fogel, Robert, 257
 Friedrich, Carl, 257
 Fudenberg, Drew, 35, 257
 fuentes históricas, 146
 Garfinkel, H., 257
 Geddes, Barbara, 142, 257
 Geertz, Clifford, 48, 50-51, 257
 Gelman, Andrew, 113, 246 n. 3, 257
 generalización, 21-22, 46-47, 53-54, 57-60, 104-5, 241
 Véase también inferencia causal; inferencia descriptiva
 George, Alexander, 56-57, 98, 178, 235, 239, 253 n. 14, 257
 Gigerenzer, Gerd, 257
 Gilpin, Robert, 257
 Goldberger, Arthur, 257
 Goldstein, Judith, 48, 202, 257
 Gould, Stephen J., 21, 257
 Granger, G. W. J., 253 n. 11, 257
 Greenhouse, Joel, 249 n. 15
 guerra, 20
 guñar y parpadear, 50-51
 Gulick, Edward, 163, 258
 hacha, ejemplo del, 17
 Hall, Peter, 202, 258
 Halpern, Nina, 203-4, 258

Hermens, F. A., 201, 258
 heurística de la disponibilidad, 252 n. 3
 hipótesis causal, 51
 hipótesis, 29
 aplicación a la recogida y análisis de datos, 23, 29-30, 39-40, 56, 57-60, 185, 203-4, 239-41
 comprobarlas con datos, 23, 30-33, 39-41, 57-60, 113-16
 concisión como objetivo de las, 30
 control de las, 40, 115
 especificidad como objetivo de las, 30
 hacerla más restrictiva, 31-32, 241
 papel de las, 20-21
 proyectos piloto al comprobarlas, 33
 reglas para construirlas, 107-24
 su falsabilidad, 30
 utilización de modelos formales para evaluarlas, 116-17
 y aumento de observaciones, 231, 234
 y endogeneidad, 198
 Hirschman, Albert O., 21, 258
 historia económica, 253 n. 7
 Hoffman, Stanley, 258
 Holland, Paul, 93, 103, 246 n. 2, 247 n. 6, 258
 homocedasticidad, 250 n. 4
 homogeneidad de las unidades, 102-5, 126-27, 241
 Hoover, Dwight, 255
 Horowitz, Donald, 94, 258
 Hsiao, C., 249 n. 1, 258
 Huth, Paul, 34, 258
 ideas como variable explicativa, 203
 identificación, 249 n. 1
 idiomas,
 conocimiento de, 137
 incertidumbre, 170, 226-29
 en la investigación científica, 19, 93
 informar de su existencia, 42, 88, 106, 163
 inclusión de variables irrelevantes, 193-96
 modelo formal de, 196

independencia condicional, 102, 105-6, 125-26
 India, 155-56, 217, 232, 239
 educación en la, 76
 y su papel en las políticas contra la pobreza, 155-56
 inferencia causal, 18, 29, 42-43, 87-124, 126-27
 con error de medida, 167
 corrección del sesgo en la, 199-200
 criterios de evaluación, 106-7, 110, 111
 incertidumbre de la, 226-29
 supuestos, 102-7
 y estudios de caso, 57
 inferencia científica, 18, 29, 42-43, 87-124, 126
 con error de medida, 167-68
 corrección del sesgo en la, 199-200
 criterios para juzgarlas, 106-7, 111
 su incertidumbre, 227-29
 supuestos de la, 102-7
 y los estudios de caso, 56
 Véase también inferencia descriptiva; inferencia
 inferencia descriptiva, 18, 26, 29, 45-57, 64, 66-85
 con error de medida, 167
 con sesgo de selección, 151
 criterios para evaluarla, 74-85
 e interpretación, 48-53, 54
 importancia en la ciencia política, 55-56
 inferencia, 57-60, 241-42
 con error de medida, 177
 corrección del sesgo en la, 199-200
 de las relaciones sistemáticas, 45-46, 91-93, 95, 247 n. 9
 eficiencia como criterio de la, 38, 76-85, 106-7, 161-62
 en la investigación científica, 18
 falta de sesgo como criterio de la, 38, 74-76, 106-7, 110, 161-62
 incertidumbre de la, 19, 88, 93, 106, 163, 170
 utilización de reglas en la, 16-17, 19, 88

Véase también inferencia causal; inferencia descriptiva
 y detalles de los hechos, 48-53, 64-66
 Inkeles, Alex, 157, 258
 instituciones, efecto de las, 26
 integración regional, 168
 interpretación, 47, 48-53, 54
 interruptor de la luz, 126
 investigación aplicada, 243 n. 4
 investigación básica, 243 n. 4
 investigación cualitativa, 15-16, 71-72 e inferencia causal, 94-96
 estilo de la, 14
 manipulación de variables explicativas, 252 n. 13
 problemas en la, 42, 55, 162-63, 241-42
 investigación cuantitativa,
 estilo de la, 14
 problemas de la, 55, 162-63, 241-42
 investigación prospectiva, 147
 investigación,
 carácter social de la, 19-20
 como procedimiento público, 18-19
 definición de, 17-20
 diseño de la, 23-24, 28-29, 128-34, 144, 185, 226-30, 240
 importancia esencial del método en la, 19
 incertidumbre, 19, 88, 93, 106
 mejorar la teoría mediante la, 29-33
 retrospectiva, 146-47, 152, 159
 ser explícito como objetivo de la, 18-19
 sobre acontecimientos complejos, 16-17, 20-23
 sobre acontecimientos únicos, 20-23
 Véase también recogida de datos; generalizaciones; investigación cualitativa; investigación cuantitativa; preguntas de una investigación
 Italia, 16, 236
 política en, 16, 236
 Iyengar, Satis, 249 n. 15 cap. 3, 249 n. 3 cap. 4, 258
 Jeffreys, Harold, 30, 258
 Jervis, Robert, 258

- Johnston, J., 258
 Jones, E. L., 54, 258
- Kahneman, Daniel, 253 n. 3, 258
 Katzenstein, Peter, 213, 258
 Keane, Mark, 253 n. 3, 258
 Kennedy, Paul, 259
 Keohane, Robert O., 48, 202, 257, 259
 Khong, Yuen Foong, 225, 259
 Kinder, Donald, 249 n. 3, 258
 King, Gary, 42, 61, 113, 140, 200, 245 n. 11, 246 n. 3, 249 n. 1, 253 n. 11, 257, 259
 Kohli, Atul, 155-56, 217, 232-33, 259
 Kreps, David, 259
 Kruger, Lorenz, 257
- Laitin, David, 157-58, 198, 217, 259
 Lakatos, Imre, 22, 248 n. 13, 259
 Lakeman, Enid, 202, 259
 Lambert, James, 202, 259
 Leamer, Edward, 248, 259
 Lebow, Richard, 258
 Leonardi, Robert, 261
 Levi, Margaret, 112, 256
 Levine, Michael, 249 n. 3
 Levy, Jack, 26, 259-60
 Lieberman, Stanley, 29, 40, 100-1, 119, 260
 Lijphart, Arend, 94, 216, 252 n. 6, 260
 Lindberg, Leon, 168, 260
 Linz, Juan, 94, 260
 Lipset, Seymour Martin, 217, 260
 Little, Daniel, 96, 247 n. 10, 260
 lógica del análisis, 16-17, 19, 88, 241
 Longino, Helen, 260
 Lowenthal, Abraham, 260
 Lynd, Robert y Helen, 37
- manchas de tinta, 31
 Mankiw, N. Gregory, 28, 260
 Mansbach, Richard, 256
 Martin, Lisa L., 15, 260
 máximo muestral, 65
 máximo, 65
 May, Earnest, 42
 McKeown, Timothy, 56-57, 239, 253 n. 14, 257
- mecanismos causales, 96-98, 238
 media muestral, 65, 77, 92-93
 media, 65, 77, 92-93
 medición de variables, 163-65
 medidas con intervalos, 163-65
 medidas nominales, 163-65
 medidas ordinales, 163-66
 Merck & Co., 260
 Merton, Robert, 18, 124, 260
 método histórico, 48-53, 54
Véase también mecanismos causales; inferencia descriptiva; investigación cualitativa
- método,
 de la diferencia, 178
 de la similitud, 178
 del compromiso, 144
 importancia esencial para la investigación científica, 19
 Middletown, 37
 Mill, John Stuart, 98, 144, 178, 252 n. 2, 260
 Miller, David, 112, 260
 Milner, Helen, 190-93, 260
 modelo de avión, 60-61
 modelos formales, 60-64, 219
 aplicados a la recogida de datos, 62-64, 116-17
 de eficiencia, 81, 111
 de endogeneidad, 208
 de ineficiencias de la variable incluída, 196
 de investigación cualitativa, 136
 de la falta de sesgo, 77, 107, 110
 de multicolinealidad, 134, 135
 del efecto causal medio, 107, 108-9
 del error de medida, 174-75, 178, 180-81
 del número de unidades, 226-27
 del problema del *n* pequeño, 130, 132, 227-28
 del sesgo de la variable omitida, 182-83, 186-88
 relevancia para la investigación cualitativa, 61
 su utilidad para las teorías, 116-17
- Moe, Terry, 260

- moneda y pluma, 119
 Moon, Donald, 48, 253 n. 13, 260
 motivaciones personales para investigar, 25
 muestreo por orden alfabético, 148
 Mugabe, Robert, 232
 multicolinealidad, 131, 133-34, 135, 158, 252 n. 4
 colinealidad relativa, 227-29
 modelo formal de la, 134-35
- nacionalsocialistas, 201-2
 Nanetti, Raffaella, 261
 Neustadt, Richard, 42, 260
 Newbold, P., 253, 257
 Nigeria, 158, 217
 notación matemática, 64-65, 68-70
 Nye, Joseph, 168, 259, 261
- O'Hear, Anthony, 113, 261
 observaciones equiparables, 106, 212-18, 225-26
 observaciones más probables, 223
 observaciones menos probables, 223
 observaciones, 62-63, 68, 127-28, 221-25
 crear muchas a partir de pocas, 230-42
 dependencia entre las, 235
 y reducción de la endogeneidad, 202-5
- Ordeshook, Peter, 116, 261
 organizaciones internacionales, 63
- Palfrey, Thomas, 249 n. 3, 261
 ¿parecerse a qué?, 216
 parpadear y guñar, 50-51
 participación voluntaria, 205-6
 Pearson, Karl, 19, 261
 persona equiparable, 212
 Plattner, Marc, 256
 Plott, Charles, 249 n. 3, 256, 261
 política comparada, área de la importancia de la descripción en, 55-56
 Popper, Karl, 24, 110, 112-14, 244 n. 5, 245 n. 7, 248 n. 13 y 15, 261
 Porter, Michael, 144-45, 261
- Porter, Theodore, 257
 postulado de sencillez de Jeffreys-Wrinch, 30
 postulado de sencillez, 244 n. 7
 precios del petróleo, 185
 precisar los componentes,
 de la variable dependiente, 200
 de la variable explicativa, 205-6
 preguntas de la investigación, 24-29
 criterios para elegirlos, 25
 su contribución a los escritos científicos, 27-28
 y estructuración de los estudios de caso, 56-57
- problema de tener más inferencias que observaciones, 129-31, 132, 136, 155, 207, 221-42
 modelo formal del, 132
 número de observaciones que se necesitan para superarlo, 226-30
 y diseño de la investigación, 226-42
 y equiparación, 216-18
 y subunidades, 233-34
 y teoría de las ciencias sociales, 240
- problema de tener más parámetros que unidades, véase problema de tener más inferencias que observaciones
- problema del archivador, 248 n. 15
 problema del *n* pequeño, 130, 132, 136, 155-56, 207, 221-42
 modelo formal del, 130, 132, 227-28
 número de observaciones necesario para superarlo, 226-30
 y diseño de la investigación, 226-42
 y equiparación, 216-18
 y subunidades, 233-34
 y teoría de las ciencias sociales, 240
- problema fundamental de la inferencia causal, 90-91, 93, 102, 105, 134, 212, 221-23
 procedimiento de la coherencia, 56-57, 251 n. 9
 pronóstico, 251 n. 7
 proteccionismo, 190-93
 proyectos piloto, 33
 Przeworski, Adam, 47, 178, 213, 216, 261
 Psathas, George, 52, 261

- público, 66
 punto de vista probabilístico, 70
 Putnam, Robert D., 16, 49, 236, 261
- Ragin, Charles, 98, 248 n. 11, 255, 261
 rastreo de procesos, 97-98, 238-41
 razonamiento analógico, 225-26
 recogida de datos, 62-63
 cuantificación en la, 55
 directrices para mejorarla, 33-36, 57-58
 fiabilidad como objetivo, 36
 garantizar la fiabilidad en la, 36, 162
 nivel de agregación, 60, 128
 para compensar la variabilidad fundamental, 227
 para maximizar el control, 216-17
 proyectos piloto antes de la, 33
 registrar el proceso de, 33-34
 reproducción como objetivo, 36-37
 utilización de modelos formales para la, 62-64, 116-17
 y eficiencia, 227, 242
 y el problema del *n* pequeño, 226-30
 régimen nazi, 201-2
 reglas de la inferencia, 16-17, 19, 88
 representación proporcional, 117-18, 164, 200, 201-2
 reproducción, como objetivo de la recogida de datos, 36-37
Resisting Protectionism, 190-93
 resumir los datos, 64-66
 estadísticos para, 64-65
 reuniones en la cumbre, 189
 revoluciones sociales europeas, 139-40
 revoluciones, 21
 Rivers, Douglas, 261
 Robinson, William, 244 n. 12, 261
 Rogowski, Ronald, 261
 Rosenau, Pauline, 261
 Rosenstone, Steven, 261, 263
 Rossi, Peter, 157, 258
 Roth, Alvin, 249 n. 3, 261
 Rubin, Donald, 246 n. 2, 261
 Russett, Bruce, 258, 261
 Ryle, Gilbert, 50
- sanciones económicas, 15
 sanciones internacionales, 15
 Sanday, Peggy, 262
 Scheingold, Stuart, 168, 260
 Schumpeter, Joseph, 17, 262
 Schwartz, R. D., 262
 Sechrest, L., 262
 Segunda Guerra Mundial, 127
 selección aleatoria, 105, 126, 242
 e investigación con *n* pequeño, 136, 155-56, 207
 sus límites, 134-38
 selección,
 de variables constantes, 157-59
 e inferencia descriptiva, 152
 en función de la variable causal clave, 157-58
 en función de la variable dependiente, 139-47, 151-59
 en función de la variable independiente, 147-49, 150-59
 intencionada, 150-59, 214
 y sesgo de la variable omitida, 214
Véase también selección aleatoria
 ser explícito, como objetivo de la investigación científica, 18-19
 series temporales, 253 n. 11
 sesgo de la variable omitida, 72-73, 100, 105, 118, 134, 178-93
 modelo formal del, 182-83, 186-88
 y asignación aleatoria de valores, 207
 y endogeneidad, 201-2
 y estudios de un solo caso, 223-35
 y selección intencionada, 214-15
 sesgo de selección, 105, 119, 127, 136-49
 e inferencia descriptiva, 152
 hacer ajustes para combatirlo, 142-44, 146-47
 inducido por el mundo, 145-47
 y fuentes históricas, 145-47
Véase también selección aleatoria
 sesgo, 38, 74-76, 106-7, 161-62, 207
 comparado con la eficiencia, 80-85
 ejemplo formal de, 77, 110
 y endogeneidad, 206
 y error de medida, 173-78
 y sesgo sustancial, 75

- Shépsle, Kenneth, 262
 Shively, W. Phillips, 26, 262
 Simon, Herbert, 262
 simplificación, 20-21, 46-47, 53-55, 57-60, 104-5, 241
 sistemas electorales, sesgo en los, 75
 Skocpol, Theda, 15, 139-40, 262
 Slovic, Paul, 252, 258
 Smoke, Richard, 257
 Snidal, Duncan, 145, 255
 Snyder, Glenn, 151, 262
 Snyder, Jack, 262
 Sober, Elliot, 30, 262
 Somalia, 157
 Stein, Janice, 258
 subunidades, 233-34
 Suppes, Patrick, 246 n. 2, 262
 supuesto del efecto constante, 103-4, 126
 Swijtink, Zeno, 257
- Tawney, R. H., 198, 262
 temas que investigar, 24-29
 criterios para elegirlos, 26
 su aportación a los escritos científicos, 27-28
 y estructuración de los estudios de caso, 56-57
- teoría de la variable oculta, 224-25, 245 n. 12, 248 n. 11
 teoría en las ciencias sociales, 29-30
 aplicación a la recogida y análisis de datos, 23-24, 29-31, 39, 57-60, 185, 240-41
 comprobarla con datos, 23-24, 30-33, 39-41, 57-60, 112-16
 concisión como objetivo de la, 30
 control de la, 40, 115
 e incremento de observaciones, 231-234
 especificidad como objetivo de la, 30
 hacerla más restrictiva, 31-33, 241
 proyectos piloto al comprobarla, 33
 reglas para elaborarla, 107-24
 su falsabilidad, 30-31
 su función, 21
 utilización de modelos formales para evaluarla, 116-17
- y endogeneidad, 198
 teorías causales, 107-8
 reglas para construirlas, 107-24
 Teune, Henry, 47, 178, 213, 216, 261
 Thursby, Jerry, 256
 Tilly, Charles, 146-47, 262
 Tirole, Jean, 35, 257
 tradiciones musulmanas en Nigeria, 158, 217
 Tversky, Amos, 252 n. 3, 258
- unidad de la ciencia, 19
 unidad de observación, 62-63, 68, 88-89
 y subunidades para aumentar las observaciones, 233-34
 utilización de subunidades para aumentar las observaciones, 233-34
- validez, 35
 validez, construir la, 35
 valores atípicos, 68
 valores esperados, 69
 variabilidad fundamental, 70, 224, 227-28, 248 n. 11
 variable causal clave, 89
 seleccionar en función de la, 157
- variables aleatorias, 62-63, 68-69, 91, 247 n. 7
 variables de control, 89
 variables de resultado, 89, 118-20
 con error de medida, 169-75
 precisar sus componentes, 200
 seleccionar en función de las, 139-47, 151-56
 su variabilidad fundamental, 227-28
 variación de las, 139-40, 144
 variables dependientes, 89, 118-20
 con error de medida, 169-75
 precisar sus componentes, 200
 seleccionar a partir de las, 139-47, 151-56
 su variabilidad fundamental, 227-28
 variación en las, 139-40, 144
 variables explicativas, 89, 133-34
 asignación de valores para las, 207-11
 con error de medida, 169-70, 173-78, 180-81

- precisar sus componentes, 205-6
seleccionar en función de, 147-49,
150-59
- variables independientes, 89, 133
con error de medida, 169, 173-78,
180-81
precisar sus componentes, 205-6
seleccionar en función de las, 147-
49, 150-60
- variables irrelevantes, 193-96
modelo formal para las, 196
- variables relevantes, 72-73, 100, 105,
118, 133-34, 178-93
modelo formal de las, 182-83, 186-88
y asignación aleatoria de valores, 207
y endogeneidad, 201-2
y estudios de un solo caso, 223-36
y selección intencionada, 214-15
- variables, 62-64, 68, 91, 247 n. 7
- variación aleatoria, 70-71
- varianza, 69-70, 92, 227-30
de la variable causal clave, 228-30
de los efectos causales, 93, 95-96,
111
ventaja competitiva, 144
ventaja electoral del que ya está en el
poder, 89-93
Verba, Sidney, 56, 178, 205-6, 233, 262
verstehen, 48-49, 52
veto, procedimientos de, 165-66
voto en Gran Bretaña, 66-69
- Walton, Douglas, 29, 263
Waltz, Kenneth, 163, 262
Webb, Eugene, 262
Weber, Max, 21, 198, 262
Weick, Karl, 262
Weiner, Myron, 75, 239, 262
Wendt, Alexander, 263
Wolfinger, Raymond, 263
Woods, John, 29, 263
- Zelditch, Morris, 157, 263
Zellner, Arnold, 30, 248 n. 13, 263
Zimbabwe, 232