

La regresión múltiple en la investigación macro-comparativa: límites y alternativas ¹

Michael Shalev

(Traducción: Mónica Fernández Fraga.
Revisión: José S. Martínez y Javier Noya)

Este artículo ofrece críticas y alternativas al uso de la regresión múltiple en la comparación internacional de los Estados de Bienestar, y en la economía política comparativa en general. Para establecer el contexto de este estudio, es bueno recordar que hubo una vez una clara división entre los estudios descriptivos y prescriptivos realizados por historiadores o analistas de políticas sociales —gente que sabía de qué estaba hablando— y los científicos sociales —gente que veía los Estados de Bienestar como una fuente de datos cómoda para comprobar generalizaciones teóricas abstractas. Esta división estaba inseparablemente ligada a las predilecciones metodológicas de cada campo. Los sociólogos y politólogos que trabajaban en el estudio de los Estados de Bienestar formaban parte de la revolución cuantitativa en los estudios comparativos. Empleando la correlación y el análisis de la regresión, con gran optimismo esperaban demostrar de forma concluyente que las dimensiones de los Estados de Bienestar son una función del desarrollo económico, la fuerza política de la izquierda y el tamaño de la población que lo necesita, o una combinación de estas explicaciones rectoras.

Este tipo de esfuerzo simplista no se ha extinguido completamente, pero ha habido una tendencia irresistible en el trabajo cuantitativo hacia una mayor sofisticación. Se han propuesto formas más complejas de análisis de regresión y algunos métodos alternativos de análisis cuantitativo. Un volumen editado en 1994 por Thomas Janoski y Alex Hicks es la fuente más destacada en este ámbito. Los comparativistas que investigan sobre el Estado de Bienestar y otras áreas de la economía política han nutrido la tendencia consolidada hacia un mayor reconocimiento de las limitaciones existentes en las aplicaciones comunes de la regresión múltiple y hacia los sofisticados intentos de superar estas limitaciones sin sacrificar el poder de la regresión. Investigadores jóvenes y con buena formación estadística, como Robert Franzese (en prensa), Torben Iversen (1998) y Bruce Western (1994, 1995) ² son brillantes ejemplos de esta conciencia crítica constructiva.

Mientras tanto, el estudio histórico-comparativo sobre los Estados de Bienestar sigue

siendo una tradición de investigación importante. Pioneros como Theda Skocpol y John Stephens, e investigadores bajo su influjo, continúan utilizando el seguimiento de procesos históricos para contrastar casos y, así, desenmarañar el rompecabezas planteado por la diversidad de Estados de Bienestar. Pero sigue en disputa la cuestión de cuál es la relación apropiada entre ese tipo de enfoques y las comparaciones cuantitativas multivariantes. Por ejemplo, mientras Skocpol evita cualquier tipo de cuantificación, Stephens y sus colaboradores creen en el diálogo constructivo entre la historia comparativa y el análisis de regresión con muchos países (Huber, Ragin, Stephens, 1991; Rueschemeyer, Huber y Stephens, 1992).

En este contexto, el libro de 1987 de Charles Ragin *The Comparative Method* representa una línea divisoria. Ragin insistía en que los estudiosos comparativistas «no deberían estar divididos en dos grupos: los que saben algo acerca de los casos empíricos reales y los que saben algo sobre las técnicas estadísticas multivariantes» (Ragin, 1987). Ragin propuso un modelo de formalización del enfoque orientado hacia los casos preferido por los analistas con una inclinación cualitativa/histórica, de manera que la ciencia social comparativa pudiese gozar del rigor de técnicas como la regresión múltiple, pero sin la carga de su epistemología «orientada hacia las variables». El Análisis Comparativo Cualitativo (ACC) se presentó como un modo concreto de desarrollar este potencial.

Tanto el giro hacia el análisis histórico-comparativo «blando» como la más «rigurosa» técnica ACC han sido objeto de amplias críticas en los últimos años. Las críticas hacen especial hincapié en las cuestiones de la falsabilidad, generalidad y reproductibilidad que supuestamente emergen cuando los científicos sociales emplean los métodos y los resultados de la investigación histórica (Goldthorpe, 1991 y 1997; Lustick, 1996). La sociología histórica comparativa ha sido reprendida por ser demasiado inductiva (Kisser y Hechter, 1991), y por no emular los principios científicos incorporados en el buen análisis cuantitativo (King, Keohane y Verba, 1994). Los críticos del ACC no están de acuerdo con que éste incorpora lo mejor tanto del análisis de casos, como del análisis de variables. En su lugar, mantienen

que sacrifica las ventajas de la explicación probabilística a la determinista, y de la teorización deductiva a la inductiva (Liebersohn, 1991 y 1994; Goldthorpe, 1997).

Un hecho llamativo de la fermentación metodológica en la política social y la economía política comparativas es la confirmación de la hegemonía de la regresión múltiple (RM). Así, Janoski y Hicks (1994) presentaban alternativas como el ACC y el «Event History Analysis», pero también se mostraban bastante interesados por la extensión de la RM a ficheros de datos de distintos países que abarcan varios momentos en el tiempo (agrupamiento o «pooling»). La antología reciente de Mjøset et al. (1997) dedicada a la controversia entre Charles Ragin y John Goldthorpe constituye una destacada contribución a la clarificación del debate, pero el tono de éste es tan general que deja intacta la tendencia común a identificar el análisis comparativo de variables con la RM. Un tercer —y todavía más evidente— indicio de esta hegemonía lo encontramos en el estudio seminal de Gøsta Esping-Andersen sobre los regímenes de Estado de Bienestar. Resulta chocante que tras ofrecer una vigorosa crítica de los supuestos esenciales del análisis de la regresión convencional, el propio Esping-Andersen recurriese de nuevo a la RM para contrastar sistemáticamente la validez empírica de sus argumentos.

Este artículo revisa y amplía las conocidas críticas de principio a la RM, pero lo hace centrándose en ejemplos prácticos que muestran el valor de los métodos alternativos. Algunos de los argumentos son de naturaleza general, pero el énfasis se pone en la investigación macrocomparativa. Surgirán dos conclusiones principales. Primero, a pesar de que hay medios técnicos para hacer frente a las limitaciones de la RM, estas soluciones o bien son poco convincentes, o bien requieren una inversión técnica tal elevada que reportan beneficios cuestionables en la actividad investigadora. Segundo, el descontento con la RM no debe llevarnos necesariamente a las alternativas cualitativas radicales como el ACC o al abandono completo de los datos numéricos. Análisis no tan técnicos (como las tablas de contingencia y los métodos gráficos) y otras técnicas estadísticas multivariantes distintas de la RM (como el análisis de conglomerados y el escalamiento multidimensional) constituyen alternativas viables y útiles.

En la ilustración más amplia de este artículo reanalizamos los datos de Esping-Andersen (1990). La RM fue una dudosa elección para los propósitos de Esping-Andersen, pero el ACC también padece serias limitaciones en esta aplicación. Por el contrario, el análisis factorial ofrece un apoyo empírico fuerte y convincente para la tesis de «los tres mundos».

1. El problema de la regresión múltiple

Dado que la RM es el buque abandonado de la investigación orientada hacia las variables en las ciencias sociales, no es una sorpresa que la mayoría de las dificultades que plantea a los comparativistas ya hayan sido enumeradas elocuentemente en el libro de Charles Ragin *The Comparative Method* (1987). Como la mayoría (aunque no la totalidad) de los métodos del análisis estadístico multivariable, la RM funciona volviendo los casos invisibles, tratándolos simplemente como la fuente de una serie de observaciones empíricas de variables dependientes e independientes. Una razón obvia por la cual esto no es deseable es que los «casos» que nos interesan en campos como la economía política comparativa son limitados en número y constituyen un universo restringido³, con lo cual son conocibles y manejables. Por lo tanto, dejar los «nombres propios» en el análisis es un modo eficaz de transmitir información y permitir a los lectores que la evalúen⁴. Es más, muchos productores y consumidores de economía política comparativa están intrínsecamente interesados en los casos concretos. Así que, ¿por qué no transigir con estos intereses manteniendo nuestros casos a la vista?

Otro motivo de debate entre los investigadores orientados hacia los casos y los orientados hacia las variables son las visiones diferentes de la causalidad. Andrew Abbott (1988) ha sostenido convincentemente que «demasiado a menudo los modelos lineales generales han llevado a una realidad lineal general, a un modo limitado de imaginar el proceso social». Abbott subraya el reducido alcance teórico de la noción de causalidad subyacente a los modelos lineales, que no permite reconocer (o es poco

probable que reconozca) situaciones en las que el efecto de cualquier variable causal dada es desigual, contradictorio (dialéctico), o forma parte de un conjunto más amplio de factores que comparten una «afinidad electiva». En el mundo social los efectos son típicamente contingentes sobre su composición concreta, en la que influyen dos tipos de circunstancia histórica: el contexto temporal (efectos de período) y el curso del tiempo (secuencias o acumulaciones históricas concretas).

El análisis orientado hacia los casos da cabida fácilmente a estas perturbaciones, porque asume desde el principio de que el efecto de cada causa depende de la amplia constelación de fuerzas en la que está engastada («causación coyuntural», en palabras de Ragin). Cuando los modelos de la RM intentan emular este supuesto, deben incluir términos de interacción que agravan el conocido «problema de los grados de libertad». La RM también es un entorno hostil para otro supuesto causal que abunda en el análisis orientado hacia los casos, a saber: que puede haber más de una constelación de causas capaces de producir el fenómeno de interés. Técnicamente, los modelos de RM pueden hacer frente a este problema añadiendo más variables independientes. Pero los resultados serán ambiguos, porque esos modelos no ofrecen un modo de distinguir las relaciones condicionales y las combinaciones diferentes de muchas causas. Sin prestar atención a los casos no podemos saber si todos aquellos con cierto valor de Y tienden a poseer alguna combinación de X1 y X2; o si la presencia de Y está asociada en algunos de ellos con X1, y en otros con X2.

Para entender completamente por qué la RM es una opción tan problemática para los comparativistas, vale la pena considerar por qué aparentemente es tan *buena* opción para ciertas clases de ciencia social. A menudo, los economistas están interesados en estimar el efecto marginal de una variable económica sobre otra, manteniendo constante el impacto de otras supuestas causas. Si los precios suben, ¿cuál será el efecto probable sobre el crecimiento económico (controladas otras influencias conocidas como la tasa de inversión o las condiciones de comercio)? Si la gente invierte en un título académico, ¿cuál será el efecto probable en sus ingresos futuros (controladas otras influencias como la raza, el género y la

experiencia laboral)? La RM se adecúa perfectamente a este proyecto. Precisamente, lo que pretende hacer es estimar los efectos marginales bajo condiciones *ceteris paribus*. Por supuesto, los economistas cuentan con la ventaja añadida de que, normalmente, tratan con variables continuas que pueden ser convenientemente tratadas empleando una métrica uniforme, el dinero. Por el contrario, mucha de la curiosidad de los economistas políticos comparativos gira en torno a la presencia o ausencia de ciertas condiciones. ¿El crecimiento económico será mayor si existen sindicatos corporativistas (o un partido socialdemócrata hegemónico, o un banco central independiente)? Estaría bien saber *cuánto* crecimiento se debe a *cuánto* corporatismo. Pero, a la hora de hacer predicciones como ésta, nuestras teorías suelen ser bastante imprecisas.

La estimación de los efectos marginales en la investigación macro-comparativa también está presidida por la ambigüedad de muchas de las variables de interés y por la dificultad de medirlas con precisión⁵. Conceptos como el de corporatismo son tan conflictivos que incluso las medidas categóricas están envueltas en un consenso dolorosamente escaso (Shalev, 1990; Janoski, 1997). Naturalmente, variables continuas, como la omnipresente «representación de partidos de izquierdas en el gobierno», se pueden medir con precisión, pero ésta en realidad oculta grandes problemas de validez y fiabilidad. Por lo tanto, las comparaciones internacionales de diferencias duraderas están plagadas de dificultades como la de que, por ejemplo, una proporción media del 50% de ministros durante cierto período es compatible tanto con un gobierno de izquierdas intermitente, como con una participación estable de la izquierda en las coaliciones de gobierno, o con un partido de izquierdas (o de derechas) dominante que ha perdido su mayoría en el gobierno. La comparación a lo largo del tiempo es también problemática, ya que los números por sí solos no pueden decirnos si el papel de la izquierda en el gobierno ha cambiado bajo situaciones cualitativamente distintas, como de partido dominante, en coalición paritaria, de miembro minoritario de una coalición, o partido bisagra frente a una derecha dividida, etc. La RM podría reflejar tal complejidad sustituyendo la medida continua de la fuerza de la izquierda por una serie de variables ficticias, o

quizás encontrando una especificación funcional no lineal que permitiese captar las discontinuidades en el efecto de la fuerza de la izquierda sobre el fenómeno de interés. Pero la primera solución desperdicia preciosos grados de libertad, y la segunda requiere o buena suerte, o un grado poco probable de sofisticación teórica.

En los subcampos de la sociología y la ciencia política, gran parte del atractivo de la RM proviene de que casa cómodamente con el método de la encuesta y el muestreo. Debido a que dispone de una proporción relativamente alta de casos por variables, la investigación con encuestas puede hacer uso de la RM como medio para introducir «controles estadísticos». A diferencia de los economistas, los investigadores pueden no estar movidos por una visión teórica del mundo que sea inherentemente marginalista. Emplean controles con la esperanza de vérselas con fuerzas causales que en un diseño experimental ideal habrían sido neutralizadas mediante la asignación aleatoria de los casos a «tratamientos» distintos. Este enfoque ha sido objeto de acalorados debates⁶. De modos diferentes, David Feedman (1991) y Stanley Lieberman (1985) han planteado argumentos convincentes según los cuales un control estadístico adecuado requeriría teorías causales mucho más sofisticadas y completas de lo que los investigadores sociales pueden esperar alcanzar. En cualquier caso, dado que los comparativistas tratan necesariamente con un pequeño número de casos, desde el punto de vista técnico es difícil que puedan analizar el efecto más que de unas pocas variables independientes.

Siguiendo con los investigadores mediante encuesta, podemos identificar una razón final de porqué el atractivo de la RM fuera de la investigación comparativa no debería inspirar su utilización en este campo. Para economizar recursos, los analistas del comportamiento electoral o de la movilidad social sólo suelen encuestar a una pequeña parte de la población que les interesa estudiar. El resultado es que, muy posiblemente, una cantidad considerable de la inmensa heterogeneidad que caracteriza a un universo como «los votantes americanos» no va a ser captada por la típica muestra de sólo uno o dos mil entrevistados. No es impensable que las combinaciones más improbables de las variables independientes —digamos, una

madre soltera perteneciente a una minoría que gana 100.000 dólares al año como alto cargo de una compañía y que vota a los comunistas—existan realmente en el universo. Desde este punto de vista, una de las ventajas de la RM es que, empleando las observaciones de las que se dispone, sus coeficientes (efectos marginales) proyectan relaciones para todas las posibilidades. Podría preocuparnos que nuestra muestra truncase artificialmente el rango verdadero, pero si esto ocurre será debido a un fallo en el diseño de investigación, y no a la elección de una técnica de análisis defectuosa.

En la investigación comparativa cuantitativa la situación es muy diferente. A menudo analizamos toda la población de estudio, y cuando no es así, la causa suele ser la falta de datos, más que las consideraciones muestrales. Por lo general, si en un conjunto de datos de varios países no se da una configuración particular de atributos, seguramente sucede que ésta no existe en absoluto. Para comprender el alcance del problema, consideremos el siguiente ejemplo hipotético, en el que empleamos sólo tres variables independientes y un nivel muy tosco de medición. El gasto público en Seguridad Social (expresado en porcentaje sobre el PIB) se «regresa» sobre el poder del partido de izquierda, sobre el grado de apertura comercial y sobre la población mayor de 65 años. Todas las variables se miden en una escala de cinco puntos. Si quisiésemos construir una tabla de entrada múltiple con esta base de datos, tendríamos 625 celdas ($5 \times 5 \times 5 \times 5$). Dado que ningún estudio de la OCDE puede tener más de 20 casos, ¡esto significaría más de 600 celdas vacías! En efecto, la RM asigna países imaginarios a algunas de estas celdas vacías para buscar el mejor ajuste lineal que se pueda generar con los datos de los que se dispone⁷. Debido a que estima los efectos paramétricos parciales como si todas las configuraciones (ajustables linealmente) fueran posibles, la RM puede dar lugar fácilmente a resultados problemáticos.

En este contexto, vale la pena recordar algunas limitaciones del modelo socialdemócrata del Estado de Bienestar (Shalev, 1983). Comparando los EE.UU. con Suecia, Andrew Martin (1973) infirió hace 25 años que la hegemonía del partido socialdemócrata era la diferencia decisiva que explicaba el compromiso post-bélico de Suecia con un Estado de Bienestar de

pleno empleo, y su notoria ausencia en los EE.UU. Numerosos estudios de correlación y regresión se hicieron eco de este argumento y pasaron a intentar confirmar su veracidad a lo largo de todo el espectro de democracias capitalistas. Surgieron dos dificultades. Una de ellas era el poder de predicción insuficiente para casos en los que no había hegemonía ni de los socialdemócratas ni de la derecha, entre otras razones por la relevancia de un complejo causal diferente (la dinámica conservadora-católica destacada más tarde por Esping-Andersen). Este es el problema de la causalidad múltiple mencionada anteriormente.

Una segunda dificultad era que el modelo socialdemócrata poco o nada podía decirnos acerca de las causas «positivas» de variación en las políticas sociales y de mercado de trabajo en los EE.UU. y en países similares. Cuando se ha comprobado mediante el análisis de regresión, los resultados han sido inferencias inverosímiles como las siguientes: en un país similar a EE.UU., pero con una década adicional de gobierno socialista, la tasa de paro sería inferior en tres puntos a la de EE.UU., y los subsidios familiares, 40% más altos. Este es un ejemplo extremo de los peligros de la generalización a partir de celdas vacías cuando cada uno de nuestros casos es un complejo *Gestalt* vertebrado históricamente. Por otro lado, una prueba de la utilidad de un modelo causal es su capacidad de dar respuesta a cuestiones contrafácticas —es decir, de rellenar celdas vacías con datos hipotéticos⁸. Así, de hecho, preguntarse cómo se habrían desarrollado las políticas norteamericanas bajo las condiciones suecas fue lo que llevó a Martín y otros autores a centrarse en la fuerza del movimiento obrero. Sin embargo, es tan improbable que algunas «celdas» se puedan rellenar alguna vez, que no deberían formar parte ni de nuestro espacio computacional, ni de nuestras predicciones. Los atributos de las sociedades no están sujetos a variaciones infinitas en combinaciones ilimitadas de unos con otros.

Desde una perspectiva de RM, este problema no es insoluble. Si una variable capaz de explicar las diferencias entre Suecia y EE.UU. no sirve de guía para comparar Canadá y EE.UU., nuestro modelo seguramente está o infra-especificado o mal especificado. Si el problema es la infra-especificación, la respuesta adecuada sería añadir variables inde-

pendientes capaces de dar cuenta de la variación en las políticas liberales. Pero si añadimos estas variables adicionales, el número total de variables puede ser demasiado grande como para ser estimado con un pequeño conjunto de datos transversales. Para dar solución a este problema, podemos estar tentados de aumentar nuestro conjunto de datos combinando observaciones de corte transversal para años diferentes. Esto tendría la ventaja añadida de que nos permitiría investigar las diferencias intra-nacionales (esto es, dentro de EE.UU.) así como entre EE.UU. y otros países. Las limitaciones de las regresiones sobre conjuntos de datos agrupados (*pooled*) se tratan en la sección 4 de este artículo.

Si el problema es una mala especificación, entonces la solución sería encontrar una explicación lo suficientemente general que pudiese dar cuenta de un abanico de variación más amplio —entre EE.UU. y Canadá, así como entre estos países y Suecia⁹. Aquí el argumento que es contrario a la RM es favorable a la disposición que tienen los investigadores orientados hacia los casos a presumir que las diferentes «familias de países» demandan trayectorias causales específicas (Castles, 1993). Si es evidente que la RM no es el mejor modo de poner a prueba explicaciones plurales, ¿cuál es? Este tema se estudiará con cierta profundidad en la sección 5, en el contexto del análisis de la tesis de Esping-Andersen sobre los tres regímenes de políticas diferentes. Se revisarán los intentos anteriores de reevaluar el sustrato empírico de esta tesis, poniendo especial énfasis en las promesas y limitaciones del ACC. También se ilustrará una alternativa tanto a la RM como a los enfoques cualitativos.

Antes de entrar en técnicas relativamente sofisticadas para tratar los problemas de «las demasiadas variables y los pocos casos» (sección 4) y de la causación plural (sección 5), discutiré dos ilustraciones concretas de las deficiencias del uso habitual de la RM. Estos ejemplos ilustrativos han sido elegidos con la intención de contrarrestar dos posibles respuestas estratégicas a la crítica general de la RM que se ha ofrecido hasta ahora. Una de ellas consistiría en moderar nuestras expectativas respecto a la RM, empleándola como un modo de dividir la varianza introducida empíricamente, más que como una estrategia para comprobar estrictamente relaciones causales

hipotéticas, y presuponiendo, por lo tanto, que no ofrece más que una aproximación vaga a la plausibilidad del modelo. La sección 2 examina un ejemplo en el que la RM se empleó en este sentido «laxo», y lo encuentra insatisfactorio. Como alternativa, uno podría sostener que, efectivamente, el estatus causal de los coeficientes de regresión debe tomarse con precaución, pero que nuestra confianza se vería reforzada si metodologías alternativas dieran lugar a hallazgos convergentes. En la sección 3 se puede encontrar una ilustración de esta estrategia de triangulación.

2. ¿«Argumentos causales» o meros «resúmenes»?

«**P**ara conjuntos de datos multidimensionales, la regresión puede ofrecernos resúmenes útiles de los datos. Sin embargo, no creo que un argumento causal pueda sustentarse mucho en una regresión» (Freedman, 1991).

David Freedman es un estadístico que cree en el poder de los números, pero cuya causa es desengañar a los científicos sociales de su creencia exagerada en la inferencia estadística como herramienta de análisis causal (Freedman, 1985 y 1991). La esencia de la argumentación de Freedman (véase también Leamer, 1983) es que el contraste estadístico de hipótesis requiere que el investigador tenga a mano una teoría muy bien desarrollada, y, a la vez, que no haya puesto la mano en los datos antes del momento en el que tiene lugar el contraste. Ya que los sociólogos no pueden o no quieren seguir estos cánones, todo lo más que pueden hacer legítimamente con la RM es emplearla para resumir conjuntos de datos multivariados.

No conozco a ningún comparativista que haya seguido este consejo al pie de la letra, pero algunos han tratado la RM como algo a medio camino entre un dispositivo formal de contraste de hipótesis y un método económico de corroborar afirmaciones empíricas amplias. Mi ejemplo de este «enfoque de pocas expectativas» proviene del estudio de Bo Rothstein (1990) sobre la variación internacional en la afiliación sindical desde una «perspectiva neoinstitucionalista». Bajo el llamado «siste-

ma Gante» los sindicatos asumen la responsabilidad de administrar los seguros de desempleo, lo que acarrea la consecuencia de que en períodos de crisis o transformación económica es poco probable que el número de sus afiliados disminuya, pudiendo incluso aumentar. Por razones teóricas Rothstein intenta demostrar que los niveles más altos de sindicalización se han alcanzado sólo en aquellos países en los que tiene lugar ese sistema. Las cifras de densidad sindical en 18 países pertenecientes a la OCDE a mediados de los años ochenta revelan que el sistema Gante está realmente presente en todos los países con los niveles más altos de penetración de los sindicatos, y sólo en esos países. A no ser que el sistema Gante esté suplantando a la estrella auténtica del «show» causal, ha sido una condición necesaria para que se den tasas de sindicación de más del 70 por ciento. Desde luego, esto no quiere decir que sea una condición *suficiente* para el superéxito de los sindicatos. Simplemente puede suceder que el sistema Gante amplifique los efectos de otras condiciones favorables.

Hay dos explicaciones de por qué una simple tabla que cruza la afiliación sindical por la presencia/ausencia del sistema Gante no es suficiente: la asociación espuria (explicaciones alternativas) y las causas adicionales (explicaciones complementarias)¹⁰. Siguiendo la convención, Rothstein pretende soslayar ambas cuestiones ejecutando una regresión múltiple que tiene en cuenta otras influencias probables en las diferencias internacionales de sindicalización. Estas variables son la participación de partidos de izquierda en el gobierno y la afiliación potencial de los sindicatos (el número absoluto de asalariados empleados y desempleados).

El modelo de Rothstein se estimó de nuevo para este artículo empleando una versión modificada de sus datos, en la que se excluye Islandia por ser de poca relevancia, y se sustituye su indicador de la fuerza de la izquierda por una alternativa más sensata¹¹. Siguiendo el original, se presentan los coeficientes *betas* estandarizados.

Densidad sindical = .47 (Gante) + .28 (gob. izq.) – .34 (logaritmo de afiliación potencial)

Todos los coeficientes son significativos a los niveles convencionales (aunque el gobierno de izquierda sólo marginalmente) y el R

cuadrado ajustado es 0,73. El coeficiente métrico para la variable Gante revela que la diferencia neta entre los sistemas Gante y no Gante alcanza un impresionante porcentaje de 27 puntos.

A pesar de estos indicios de éxito, se puede argumentar que el empleo que hace Rothstein de la RM no es el más apropiado, e incluso que es, en parte, erróneo. Rothstein es comedido en sus palabras al decir «que las tres variables tienen un efecto explicativo independiente de aproximadamente el mismo tamaño estandarizado» (Rothstein, 1990). Un requisito previo para poder tomarse en serio estos «efectos explicativos» es que el modelo sea plausible desde el punto de vista teórico. Sin embargo, un poco más arriba el propio Rothstein suscita dudas sobre esto. Así, tacha el argumento sobre el significado del número potencial de afiliados de lógicamente insostenible, y el argumento del gobierno de izquierda, de viciado por lo que los estadísticos llamarían sesgo de simultaneidad. Además, aunque los coeficientes estandarizados sugieren efectivamente que el sistema Gante tiene mucho más peso empírico que las explicaciones alternativas, como los países son invisibles, los resultados no sustentan la principal afirmación de Rothstein de que es el sistema Gante, y no la fuerza de la izquierda o el tamaño potencial pequeño, lo que diferencia los países con mayor afiliación del resto de los países. Ciertamente, habría que desechar por defecto esta afirmación si el efecto del sistema Gante hubiese desaparecido una vez que se añadiesen a la ecuación las otras variables. Pero la regresión no puede aportar pruebas positivas a favor de los argumentos de Rothstein.

Más allá de estas limitaciones específicas de la RM en el argumento de Rothstein, hay que hacer notar que la especificación de su modelo descansa sobre supuestos típicos, pero cuestionables. El primero es el habitual de que las variables explicativas ejercen efectos independientes, antes que interconectados, lo que obviamente es dudoso en éste y en la mayoría de los modelos comparativos de países. El segundo es no suponer que ninguno de los efectos está condicionado por el valor de otras variables –es decir, no considerar la posibilidad de que haya efectos de interacción.

Un modo sencillo de encarar estos problemas es resumir causas y efectos en un formato de tabla descriptiva que identifique diferentes

Figura 1

| | | Países grandes (máx. 5 millones) (Media = 30) | | | | Países pequeños (máx. 10 millones) (Media = 62) | | | | |
|--|----------|--|--------------|----------------------------------|---------|--|-----------|----------------------------|-----------------------------|----|
| | | Sin izquierda (Media = 29) | | Izq. media/débil (Media = 31) | | Izq. débil (Media = 51) | | Izq. media (Media = 58) | Izq. Fuerte (Media = 71) | |
| Sin sistema Gante (Media = 0,39) | EE.UU. | 18 | Francia | 15 | Suiza | 34 | Holanda | 29 | Austria | 57 |
| | Japón | 28 | Italia | 36 | Irlanda | 68 | Australia | 51 | Noruega | 58 |
| Con sistema Gante (Media = 0,81) | Alemania | 31 | Gran Bretaña | 43 | | | Bélgica | 74 | Dinamarca | 83 |
| | Canadá | 38 | | | | | Finlandia | 80 | Suecia | 86 |

combinaciones de condiciones (causas) con los países que las «encarnan». Esto requiere algo de previsión, porque el modelo de Rothstein se refiere a tres variables causales diferentes, y su variable dependiente, la afiliación sindical, no es fácil de agrupar (su distribución es bastante homogénea a lo largo de un amplio rango de valores). La solución que proponemos es un simple «árbol», o diagrama de flujos de varias entradas, que muestra los valores exactos de afiliación sindical para diferentes «conglomerados» de países. Estos grupos se crearon de manera muy sencilla, cruzando en una tabla la presencia o la ausencia del sistema Gante por versiones categóricas de las otras dos variables causales de Rothstein ¹².

Los resultados (figura 1) ofrecen pruebas interesantes de la existencia de efectos causales anidados. Se deducen fácilmente, en primer lugar, por la diferencia sistemática entre las configuraciones existentes y las no existentes. Sólo en los países pequeños se alcanzó una representación considerable de los partidos de izquierdas, y sólo los países con una izquierda fuerte muestran el sistema Gante ¹³. En el caso de la afinidad entre el sistema Gante y la fuerza de la izquierda, no podemos decir en qué sentido apuntan las flechas de la causalidad sin embarcarnos antes en la investigación histórica. (En realidad, lo mismo se puede decir de la

relación entre la afiliación, por un lado, y la fuerza de la izquierda y el sistema Gante, por el otro). Pero se puede colegir que lo que va asociado con niveles más altos de sindicación es un *síndrome* de pequeño tamaño, «izquierdismo» y sistema Gante. En los resultados también se observan indicios de una interacción. El «efecto Gante» parece mayor en el caso de los países en los que la fuerza de la izquierda es moderada que en el de las social-democracias consolidadas.

Este método tan «poco sofisticado» de presentar los datos hace justicia a la tesis de Rothstein con mucha más eficacia que su análisis de regresión. Aclara precisamente aquello que quería demostrar, a saber: que el efecto Gante es importante, y no espurio, y que entra en acción en países en los que otras condiciones son ampliamente favorables a los sindicatos. Pero estos resultados hacen otra cosa importante: indicar al investigador interesado las cuestiones más fructíferas que, en las comparaciones selectivas de casos, pueden ayudarlo a establecer la importancia real del sistema Gante ¹⁴. En concreto, debemos cuestionar si el sistema Gante por sí solo puede explicar las grandes diferencias de densidad entre los miembros de dos parejas de países, por lo demás muy semejantes: Bélgica y Holanda, por un lado, y Suecia y Noruega, por otro ¹⁵.

La visibilidad de la relación entre variables y casos, en la sencilla presentación mediante diagramas que aquí defendemos, puede ponernos en la pista de casos anómalos que revelan las limitaciones del modelo teórico. Prestar atención a las observaciones atípicas (*outliers*) del análisis de la regresión también es, a veces, un modo de identificar anomalías, pero no del tipo de las estudiadas aquí, es decir, países que no «tienen sentido» cuando se ponen en relación con otros casos similares. La presentación de los datos en tablas o en gráficos con las observaciones mencionadas sí permite esto; la inspección y el diagnóstico de los residuos de la regresión, por contra, no lo permite.

3. Crear confianza mediante la triangulación

Peter Hall y Robert Franzese (1998) son los autores de un estudio excepcionalmente minucioso que cuestiona uno de los dogmas más influyentes de la economía política moderna, a saber: que los bancos centrales independientes tienen un efecto favorable muy intenso sobre los resultados económicos. Sostienen que, si bien los bancos independientes son siempre anti-inflacionistas, también existe el riesgo de que su impacto sobre el desempleo sea mucho menos saludable. A no ser que la fijación de salarios sea centralizada y coordinada, los negociadores fracasarán en la internalización de las «señales» bancarias, y el resultado será una tasa de desempleo más bien alta que baja.

Para contrastar su argumento, Hall y Franzese siguieron en tres pasos. En primer lugar, demostraron su plausibilidad tomando el caso paradigmático de Alemania Occidental. En segundo lugar, usaron datos de 18 países de la OCDE a lo largo del período de posguerra (1955-1990) presentándolos en un formato de tabular simplificado. Finalmente, emplearon la RM para poner a prueba un modelo más elaborado en varios niveles de «agregación», abarcando desde medias para períodos completos de tiempo (puro corte transversal) hasta datos anuales «pooled». Los resultados de cada uno de estos análisis son coherentes con su argumentación de que el impacto del status del banco central viene condicionado por la estructura de la negociación salarial.

En su análisis tabular Hall y Franzese agrupan las medidas tanto de la independencia del banco central como de la coordinación de salarios, y las cruzan, primero en forma de una escala de tres puntos y después como dicotomías. Los resultados confirman claramente la hipótesis del efecto de interacción. Sin embargo, este efecto podía ser en principio un artefacto, el resultado de alguna influencia confundente como la riqueza de un país, la apertura económica o la composición del gobierno. Pero el resultado inicial sobrevive a la aplicación de la RM empleando controles para las variables económicas, políticas e institucionales clave. Las estimaciones condicionales de los parámetros muestran que la interacción entre la independencia y la coordinación es significativa desde el punto de vista sustantivo y estadístico. Y el contraste de diagnóstico de la regresión muestra que estos resultados no dependen de la presencia de ningún caso particular.

Si bien los autores conducen su análisis de regresión con profesionalidad, es heroico imaginar que con 18 observaciones se puedan descifrar los efectos de un total de nueve variables «independientes» (incluyendo los términos de interacción). En este punto Hall y Franzese parecen poner los pies en la tierra, porque (a) realizan algunas de sus regresiones empleando datos agrupados (*pooled*) que obvian el problema de los grados de libertad, y (b) las regresiones se construyen sobre los resultados bien definidos (aún «sin controlar») de su análisis de tablas.

El único elemento que se echa en falta en los resultados de Hall y Franzese es el mismo que resultó ser crucial en nuestro re-análisis del estudio de Rothstein: la identificación de los países. Además, como en cualquier análisis de datos recodificados, el resultado depende de decisiones acerca de cómo agregar los datos brutos. Se puede emplear una gran variedad de criterios: conocimiento sustantivo de los casos, agregación en categorías de tamaño similar, o ajuste de las nuevas categorías a los «cortes» en la distribución de observaciones. La inspección de los datos de partida sobre la coordinación no sugiere ninguna base «natural» para agrupar cualquiera de las cinco categorías originales. Los datos sobre la independencia bancaria forman una distribución continua pero asimétrica. Tras inspeccionar visualmente las puntuaciones brutas, habiendo encontrado dos rupturas aparentes, los convertí en puntuaciones bajas (<.4),

Tabla 1. Reanálisis de Hall y Franzese *

| COORDINACIÓN DE LA NEGOCIACIÓN SALARIAL | | | | | | |
|---|-----------------------|----------------------|--------------------|---------------------------------|-------------------|-------|
| | 0 | 0.25 | 0.5 | 0.75 | 1 | n |
| Independencia del Banco Central | | | | | | |
| Baja (<.4) | | ITALIA N. ZELANDA | | | NORUEGA SUECIA | 4 |
| Media | IRLANDA INGLATERRA | AUSTRALIA FRANCIA | BÉLGICA HOLANDA | DINAMARCA FINLANDIA JAPÓN | | 9 |
| Alta (>.6) | CANADÁ EE.UU. | | | SUIZA ALEMANIA | AUSTRIA | 5 |
| n | 4 | 4 | 2 | 5 | 3 | |
| COORDINACIÓN DE LA NEGOCIACIÓN SALARIAL | | | | | | |
| | 0 | 0.25 | 0.5 | 0.75 | 1 | Media |
| Independencia del Banco Central | | | | | | |
| Baja (<.4) | | (1.3, 5.6) | | | 2.0 | 2.7 |
| Media | (4.9, 8.1) | 4.1 | 4.9 | 93.3 | | 4.5 |
| Alta (>.6) | 6.1 | | | (.9, 3.1) | 2.2 | 3.7 |
| Media | 6.3 | 3.8 | 4.9 | 2.8 | 2.0 | 3.9 |

* La parte de arriba muestra los países; la parte de abajo, las tasas de desempleo.

medias y altas (>.6). El primer panel de la tabla 1 muestra los países que ocupan cada celda, y el segundo, el nivel de desempleo.

El primer panel de la tabla revela que 6 de las 15 celdas posibles están vacías. Afortunadamente, los huecos no están demasiado juntos, pero las celdas vacías necesariamente hacen más difícil inferir relaciones continuas. Dicha inferencia también se ve complicada por otro rasgo de los datos: en tres de las celdas ocupadas por un par de países, estos tenían tasas de desempleo ampliamente divergentes (se muestra por separado).

Los totales marginales del panel inferior son sólo parcialmente coherentes con las expectativas de Hall y Franzese. La fila inferior muestra que el desempleo tiende a caer a medida que aumenta la coordinación, aunque hay unas cuantas excepciones¹⁶. Pero la columna de la derecha muestra que la tasa media de desempleo más alta se da en países con un nivel medio de independencia bancaria, más que en

países de niveles altos. La inspección individual de las columnas sugiere que la independencia bancaria supone poca o ninguna diferencia sea cual sea el nivel de coordinación. Sin embargo, se intuye una interacción moderada cuando observamos las filas individualmente: la tendencia de la coordinación a reducir el desempleo es muy clara entre los cinco países con los bancos centrales más autónomos.

Por supuesto, esta debilidad de las pruebas a favor de la hipótesis de que el efecto bancario está condicionado por las instituciones de la negociación salarial puede ser el resultado de ignorar las influencias teóricamente extrañas que Hall y Franzese controlan en sus regresiones. Desde luego, dadas las críticas a la RM expuestas más arriba en este artículo, me resistiría a aceptar las evidencias dadas por el análisis de la RM en sentido literal. Un enfoque alternativo es volver al panel superior de la Tabla 1, donde se nombran los 18 países. La expectativa del autor de que, en ausencia de

una negociación salarial coordinada, la independencia del banco central acentuaría el dilema desempleo-inflación, puede explicar la gran diferencia en la tasa de paro entre las «coordinadas» Austria, Suiza y Alemania, por un lado, y los dos países «no coordinados» de América del Norte, por otro. Hall y Franzese necesitan persuadirnos de que es realmente la coordinación, y no factores como la apertura económica o la representación política del trabajo, lo que explica este hiato en particular.

Desde el momento en que se especifica este contraste, inmediatamente vienen a la mente dos reservas. En primer lugar, cada uno de estos subgrupos de países constituye una región cultural, política y económica que tienen numerosas similitudes internas, pero que son mundos aparte. El supuesto característico del AR, de independencia de las observaciones, se antoja especialmente anómalo en un contexto como éste, en el que emergen estos conglomerados que comparten tantas características. En segundo lugar, sabemos que en la década de los noventa estas dos subregiones sufrieron cambios opuestos en sus respectivos resultados pasados referidos al desempleo. Replicar el análisis empleando datos más recientes nos proporcionaría un test evidente de la robustez del modelo. A menos que los cambios en las variables de control hayan creado disparidades de magnitud similar, es cuestionable que este nuevo contraste fuese favorable.

En resumen, el peso del análisis comparativo de Hall y Franzese descansa en sus regresiones múltiples. Pero, en tanto en cuanto la validez de su argumentación se basa en cinco casos críticos, los autores deberían reconocer esto explícitamente llevando a cabo comparaciones razonadas y sistemáticas de países (con su nombre) y/o períodos. Incluso un simple análisis de tablas que no facilite los nombres delega la tarea de la generalización a una «caja negra» (en ese caso, valores medios basados en casos anónimos). Por consiguiente, no se puede considerar que la triangulación está completa si no pasamos al nivel de los casos, en el que el conocimiento por parte del investigador de los casos individuales, junto con el empleo creativo de comparaciones selectivas de casos, puede desempeñar un papel decisivo a la hora de apuntalar o cuestionar la plausibilidad de generalizaciones de nivel superior.

4. ¿Es el agrupamiento de datos (*pooling*) una panacea?

Algunos lectores pueden ver en mis comentarios a los dos artículos revisados hasta ahora otra ilustración más de un problema bien conocido: como los comparativistas tienen «demasiadas variables a la caza de pocos casos», la RM sólo se puede aplicar toscamente (Rothstein) o de forma poco plausible (Hall y Franzese) en diseños transversales estándar. Mi enfoque alternativo puede criticarse por ser una retirada deshonrosa hacia una interpretación de resúmenes descriptivos de datos que depende demasiado de decisiones arbitrarias sobre cómo agruparlos y presentarlos. Sin lugar a dudas, estos críticos rechazarían mi argumento de que, básicamente, la regresión es inadecuada para el análisis macro-comparativo, y preferirían concentrar directamente sus energías creativas en la solución del problema de la insuficiencia de casos (Cf. King, Keohane y Verba, 1994).

En sintonía con esta segunda postura, John Goldthorpe ha defendido que «en el fondo el problema del N pequeño no es de método, sino de datos». En concreto, Goldthorpe recomienda emular al elevado número de investigadores que han agrupado («pooled») datos del mismo conjunto de países para diferentes épocas. De esta manera, «las observaciones –y los grados de libertad– aumentaron...»¹⁷. Sin embargo, hay razones muy sólidas para creer que la consecuencia más probable de un giro hacia el agrupamiento será enturbiar todavía más las aguas causales.

¿Qué acarrea el agrupamiento?¹⁸ Tradicionalmente la investigación cuantitativa en el nivel macro analizaba o bien instantáneas de países diferentes en un momento concreto (datos de corte transversal), o bien datos de un único país referidos a épocas diferentes (cifras anuales o «paneles» sub-periódicos). Por el contrario, los ficheros de datos agrupados apilan series temporales o datos de panel para varios países unos encima de otros. De esta manera, incorporan tanto las variaciones comparativas entre países, como la variación dinámica a lo largo del tiempo. Como resultado, los analistas deben bregar con las complicaciones

técnicas características *tanto* de la estimación transversal como de las series temporales. No es de extrañar que los investigadores se enfrenten a un desconcertante abanico de opciones técnicas y escollos, modas y manías muy volátiles, así como también a consejos contradictorios¹⁹. No entraré en estos problemas aquí. Más bien, mi interés se hace eco de la conclusión de la guía más amplia al respecto, en el sentido de que el agrupamiento de datos puede ser contraproducente «si de antemano no se sopesa seriamente el *significado* de cada una de las agregaciones al «pool» de datos» (Sayrs, 1989).

Muchos de los investigadores comparativistas que «agrupan» lo hacen movidos por la misma agenda que inspira el uso de regresiones de corte transversal, es decir, el interés en explicar las diferencias duraderas entre países. Para estos investigadores, el agrupamiento es esencialmente un modo de multiplicar el número de «instantáneas» transversales a su disposición. Están deseosos de tratar cada «instantánea» como una visión más de la variabilidad entre países. Como consecuencia del incremento considerable en el número de observaciones en un conjunto de datos agrupados, esperan superar las serias limitaciones del análisis causal multivariable que resultan del reducido número de países que forman el universo de los Estados de la OCDE.

Por otra parte, los datos agrupados también son de utilidad para los investigadores interesados en explicar las *diferencias nacionales en procesos dinámicos*. Por ejemplo, hace dos décadas los investigadores (incluido el que escribe) estuvieron interesados en comparar las causas de las fluctuaciones en la actividad huelguista en diferentes países (Hibbs, 1976; Shalev, 1979). Algunos investigadores consideraron que los resultados divergentes obtenidos de la aplicación de las mismas regresiones de series temporales en diferentes países eran simples antidotos contra las generalizaciones excesivas (Paldam, 1982). Pero otros vieron en estas divergencias ejemplos del efecto predecible de las variaciones contextuales (Zinder, 1975). Como explico más abajo, éste ha sido el rumbo seguido por los analistas más concienzudos de los ficheros de datos agrupados, Larry Griffin, Larry Isaac y sus colaboradores (Griffin, 1986; Griffin, 1989).

Una segunda dificultad *a priori* en los diseños con datos agrupados radica en que el efec-

to de una variable independiente dada puede ser muy diferente en las series temporales y en los datos de corte transversal «porque las estructuras causales subyacentes difieran» (Firebaugh, 1980). Walter Korpi y yo observamos justamente esta divergencia en nuestro estudio de las huelgas. Mientras que las fluctuaciones temporales en las huelgas seguían una lógica económica, en la que una menor tasa de paro estimulaba la militancia obrera, la variación transversal seguía una lógica político-económica, en la que una menor tasa de paro operaba como *desincentivo* al empleo del arma de la huelga entre los movimientos obreros fuertes (Korpi, 1980). En consonancia, Alex Hicks (1994) ha instado a aquellos que agrupan los datos a emprender «comparaciones sistemáticas de fuerzas causales que varían transversal y longitudinalmente»²⁰.

A pesar de la existencia de estas razones para no tratar los datos agrupados de manera simple, demasiado a menudo el análisis multivariable de los datos agrupados procede como si los datos viniesen dados por un único conjunto de determinaciones invariables. Pero las consecuencias prácticas de este supuesto son bien conocidas. En la más temprana y, por el momento, mejor explicación del agrupamiento de datos, Griffin et al. (1986) usaron los datos anuales de 12 países y 16 años para investigar los efectos de seis variables económicas y políticas sobre el gasto público destinado en esos países a ayudas sociales. Su primer hallazgo fue que el grueso de la variación en la mayor parte de sus variables independientes se concentraba o bien en la dimensión del tiempo, o bien en la de los países. Ya sólo esto nos está diciendo que no habría tenido sentido emplear un único modelo para explicar ambas dimensiones. Y, de hecho, Griffin et al. encontraron que «la pendiente media entre países y la pendiente media de las series temporales... tenían muy poco en común» (p. 116). Incluso dentro de cada una de las dimensiones, la del tiempo o la del espacio, no se puede ignorar la contextualidad de las relaciones causales. Los resultados de las secciones transversales anuales resultaron ser «extraordinariamente inestables a lo largo de los años», incluso cuando se trataba de años consecutivos (p. 111). Aunque las estimaciones de las series temporales específicas de cada país fueron más estables, parecían «evocar procesos (nacionales) muy dife-

rentes subyacentes a las transferencias sociales» (p. 115).

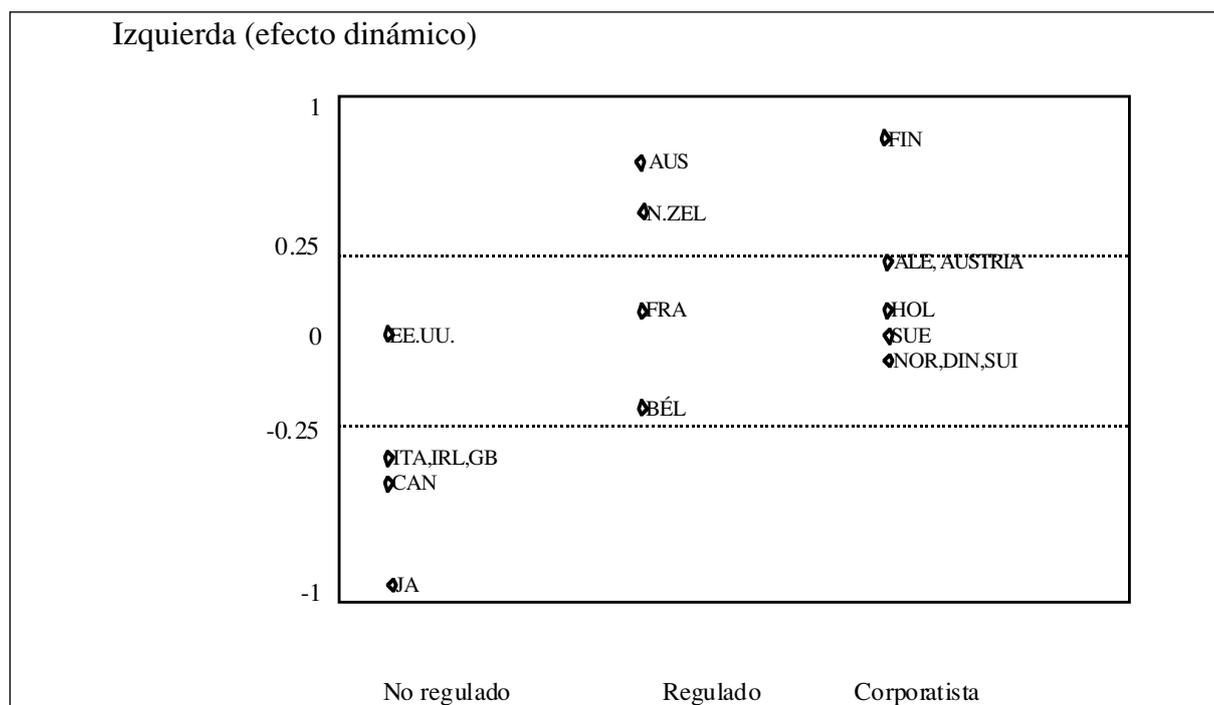
Quizás no suceda siempre que todas las bases de datos agrupados están plagadas de discontinuidades tan graves como las descubiertas por Griffin et al. Pero los investigadores que emplean secciones transversales múltiples, con la esperanza de comprar más grados de libertad, cuando menos, deberían comenzar su análisis confirmando empíricamente las uniformidades temporales y espaciales que deben tomar como supuestos de partida ²¹. De todos modos, creo que tiene más sentido reconocer que las diferencias a largo plazo entre los países, y las variaciones a corto plazo en el tiempo, plantean cuestiones diferentes y, potencialmente, invitan a diferentes explicaciones. Esto no significa necesariamente que haya que desterrar los usos enfocados y «responsables» del agrupamiento de datos. Aún cuando nuestra preocupación principal sea explicar la variabilidad nacional, podemos volver a los datos agrupados para ver si los patrones de causalidad varían a lo largo del tiempo, o por el contrario, aislar momentos decisivos de cambio en los que las diferencias entre países quedaron «fijadas» por primera vez, o mudaron marcadamente. Por otro lado, para quienes están interesados principalmente en explicar los procesos dinámicos, el agrupamiento de datos hace posible contemplar múltiples explicaciones adaptadas a diferentes contextos. La dinámica característica de un país (o un grupo de países) puede verse tanto como una indicación o como una causa de diferencias (estructurales) duraderas.

Griffin y sus colegas propusieron a los investigadores interesados en este tipo de problemas una metodología sistemática. Sugirieron estimar los parámetros de las series temporales en regresiones para cada país por separado. Después, en una segunda vuelta, estos parámetros podían tratarse como variables dependientes a explicar transversalmente por las diferencias entre países (Griffin, 1996). Aunque la técnica daba resultados sugerentes (Griffin, 1989), la credibilidad de estos resultados depende de nuestra confianza en las estimaciones de las series temporales de la primera vuelta. Dado el escaso número de observaciones implicadas –y, para nada más que un corto período, deberíamos sospechar que existe inestabilidad causal– resulta difícil no ser escéptico.

Recientemente Bruce Western (1996) ha sugerido cómo puede superarse esta limitación. Western está especialmente interesado en explicar la dinámica del desempleo. Quiere demostrar que factores institucionales, como la presencia o la ausencia de corporatismo, pueden explicar las diferencias entre países en los efectos de las fluctuaciones de variables como la composición del gobierno o el gasto social. Reuniendo una serie de datos agrupados para 18 países de la OCDE entre 1964 y 1990, Western selecciona deliberadamente una mezcla de variables explicativas «dinámicas» e «institucionales». Aboga por un enfoque bayesiano de estimación que permita posibles diferencias contextuales en la dinámica causal, pero que se diferencia del método de las dos etapas de Griffin en un aspecto importante. La técnica de Western posibilita que las estimaciones para países individuales «absorban» de toda la muestra. Las implicaciones de esta estrategia son muy profundas. Parece permitir que el analista saque partido de las numerosas observaciones y la gran diversidad proporcionadas por los conjuntos de datos agrupados, sin tener que presumir una causalidad idéntica en el tiempo y el espacio ²². Entonces el agrupamiento se liberaría, entonces, de la mayoría de las objeciones que he presentado y, como explica Western, la cuestión de si los comparativistas deberían generalizar dentro, o más allá, de los contextos específicos se convertiría en una cuestión que puede abordarse empíricamente, dejando de ser un problema epistemológico irresoluble.

A este respecto, el éxito de Western se valora mejor si se consideran los resultados de su análisis ilustrativo del desempleo. Es impresionante como llega a demostrar las implicaciones del corporatismo a largo y a corto plazo. A largo plazo, los países corporatistas experimentan tasas de paro significativamente inferiores a las de otros países. Al mismo tiempo, en los países corporatistas el impacto dinámico del crecimiento salarial sobre el desempleo también es bajo, lo que confirma la afirmación de que el corporatismo salvaguarda el empleo al mejorar a corto plazo el dilema entre salarios y empleos. A pesar de todo, una vez más, es necesario comprobar la credibilidad de las conclusiones estadísticas revisando los casos individuales, y en este ejemplo tampoco el resultado es alentador. Consideremos

Figura 2. Modelo jerárquico del desempleo, de Western



los coeficientes de las series temporales para cada país (Tabla 4 de Western). En los países escandinavos (que suponen la mitad del grupo corporatista) encontramos efectos *positivos* del crecimiento salarial sobre el desempleo²³; y –lo que resulta igualmente extraño–, dentro de los países con coeficientes negativos, la mitad *no* son corporatistas.

Los hallazgos de Western sobre los efectos dinámicos de los cambios en la composición del gobierno son todavía más confusivos. Parecen demostrar que en los países corporatistas, y en otros escenarios en los que la negociación colectiva está muy extendida, *los incrementos en el poder del partido de la izquierda tienen como resultado el crecimiento del desempleo*. Para llegar al fondo de este resultado contraintuitivo, la figura 2 reproduce para cada país las estimaciones de Western sobre el efecto dinámico de los cambios en la representación de la izquierda en el gobierno. Con el fin de destacar los posibles efectos institucionales del tipo que interesan a Western, los países se han agrupado en varios contextos con alguna semejanza. (Los mercados laborales «no regulados» son aquellos en los que no más de la mitad de la fuerza de trabajo está cubierta por

acuerdos colectivos. Nótese que, de acuerdo con este criterio, todos los países corporatistas, excepto Suiza, son «regulados»).

A primera vista, la figura parece confirmar el hallazgo de que en los mercados de trabajo corporatistas/regulados «los gobiernos socialdemócratas tienen como resultado un aumento del desempleo» (Western, 1996). Sin embargo, también es evidente que, sin dos casos atípicos –Japón y Finlandia–, esta tendencia no sería significativa. El problema es que para ninguno de estos dos países resultan plausibles los coeficientes dinámicos de Western. Finlandia experimentó escasos cambios destacados en el papel general de la izquierda en el gobierno (tenía uno de los coeficientes de variación más bajos, de sólo 0,23). Lo que sí *varió* en Finlandia fue el papel relativo de los partidos comunistas y socialistas, pero, por supuesto, la medida empleada por Western obviaba esto. Japón es un candidato todavía peor para ejemplificar tendencias, ya que, en el período relevante para el estudio, su representación del partido de la izquierda fue un cero permanente²⁴.

Me inclino a sacar la conclusión de que el sueño de, simultáneamente, aprovechar la abundancia de información de las bases de

datos agrupados, al tiempo que se respeta e, incluso, se explora la diferencia entre la causalidad sincrónica y diacrónica, sigue siendo inalcanzable. Se recordará que el mecanismo que supuestamente reconciliaba estos dos objetivos era descrito como «fuerza prestada». En palabras de Western (1996), «la información de otros países ayudará a obtener una estimación de un coeficiente para un país concreto en el que, digamos, una determinada variable independiente no muestra variación». Este procedimiento sería una versión más del tema estándar en la regresión –el referido a la inferencia de relaciones causales sobre datos inexistentes– si no fuese por una cosa. En el modelo jerárquico de Western, los vínculos dinámicos inexistentes se convierten en la materia prima para la contrastación real de los efectos institucionales. A nadie le puede sorprender que un análisis que utiliza datos imaginarios como «input» corra el riesgo de generar un «output» inverosímil.

Es una pena, porque la premisa estratégica de Western era sugerente. Nuestras vidas serían más fáciles si se pudiese solventar en el terreno empírico la cuestión de si la causalidad es contextual –con lo cual, son indispensables los nombres propios– o general –los nombres propios capitulan ante los nombres de las variables–. Pero, por supuesto, una empresa tal obliga a que nuestros conceptos se puedan operacionalizar de modo que permitan la generalización. Uno de los requisitos es que haya suficiente variabilidad. Para poner en cuestión lo obvio una vez más, si nunca ha habido –y puede que nunca haya– *ningún* socialista en el congreso, la representación socialista en el gobierno es de poca utilidad para dar cuenta de un *explanandum* que incluya las fluctuaciones del desempleo o del gasto social en América. Incluso aunque haya suficiente variabilidad, como han subrayado Przeworski y Teune, necesitamos buenas medidas, es decir, medidas *comparables*. La proporción de ministros de partidos de izquierda en el gobierno es una medida que ejemplifica los problemas típicos que se le plantean a los comparativistas. Por un lado, se puede medir uniformemente, y bastante fácilmente, en los distintos países. Por otro lado, un cambio de un simple punto porcentual en la representación de la izquierda de un año (u otro subperíodo de tiempo) para otro, puede significar algo muy diferente según cuál sea el

país y el período. Estas diferencias pueden traducirse perfectamente en múltiples conceptos generalizados, pero en la práctica esto nos llevaría de nuevo al problema de los grados de libertad. Una vez más, observamos que, realizado correctamente, el agrupamiento de datos no puede eludir este problema.

5. El enfoque de los regímenes, a prueba

La obra de Gøsta Esping-Andersen *Los tres mundos del Estado del Bienestar* suscita una cuestión profunda a los comparativistas que, como él, están de acuerdo en la necesidad de (1) reconocer que puede haber discontinuidades causales sorprendentes entre contextos diferentes; (2) sustanciar sus hipótesis sobre las relaciones entre las variables, haciendo uso de su conocimiento de los casos; y (3) emplear indicadores cuantitativos para contrastar de manera sistemática proposiciones en todo el universo de casos. Al igual que en los dos primeros estudios revisados en este artículo, Esping-Andersen adoptó un enfoque de dos etapas. Desarrolló índices de «universalismo», «desmercantilización» y «estratificación», y usó tablas para demostrar que sus 18 países de la OCDE tendían a diferenciarse en tres subgrupos distintos (Esping-Andersen, 1990: Tablas 2.1, 3.3, 4.3). Después, utilizó la RM para realizar un análisis causal de la variación entre países de más de una docena de indicadores, algunos brutos y otros agregados, que se «regresaron» sobre variables políticas y, también, en algunos casos, demográficas y sobre otras de control (Esping-Andersen, 1990). Soy de la opinión de que la primera técnica de Esping-Andersen era innecesariamente «blanda», mientras que la segunda está explícitamente reñida con su postura epistemológica. Tras aclarar estas afirmaciones, propondré una estrategia analítica alternativa.

El análisis de tablas de Esping-Andersen se basaba casi exclusivamente en la intuición del autor –tanto en lo relativo a la construcción de índices agregados como a la identificación de conglomerados de países. No se realizó ningún contraste *sistemático* para comprobar si, realmente, el *conjunto* de indicadores de los regí-

menes de bienestar «tenía algo que ver» entre sí; y, de ser así, si los países se agrupaban realmente en tres subgrupos distintos en dimensión(es) específicas de las políticas públicas. Habría sido un paso lógico someter estas afirmaciones a un análisis con técnicas como el análisis factorial, el análisis de conglomerados o el escalamiento multidimensional.

Como hemos dicho, el giro de Esping-Andersen hacia la RM tenía la intención de fundamentar sus afirmaciones causales. Uno de los modos en que intentó hacer esto fue demostrando que sus variables (políticas) preferidas obtenían un apoyo empírico mayor que el de las variables explicativas alternativas (p.ej., las demográficas). Sin embargo, la credibilidad de este ejercicio era limitada, dado que implicaba hacer regresiones con 5 o 6 variables «independientes», pero sólo para 18 casos. El propósito central del análisis de regresión era refrendar empíricamente las hipótesis relativas a los antecedentes políticos de los regímenes de políticas públicas. Pero este enfoque implicaba un desajuste obvio entre las afirmaciones de Esping-Andersen y sus métodos. El argumento causal clave en el libro es que *los países se agrupan en torno a políticas públicas porque se agrupan en torno a la política*. Sin embargo, el enfoque de la regresión presupone que tanto las políticas públicas como la política son variables continuas, y no un número limitado de configuraciones cualitativamente diferentes con raíces históricas distintas.

Queda de manifiesto la incompatibilidad fundamental entre la RM y el enfoque de los regímenes de Esping-Andersen. Resulta difícil pensar que pueda afirmar que, digamos, cualquier incremento discreto del catolicismo o del absolutismo debería producir un incremento discreto y uniforme del «corporatismo» en las pensiones. *Sólo* de los países que son mayoritariamente católicos, y/o tienen un pasado absolutista, se espera que muestren un perfil de políticas corporatistas. Del mismo modo, él no diría que las políticas sociales de un país determinado pueden entenderse precisamente como el efecto combinado del catolicismo, el absolutismo y la movilización de la clase trabajadora (como en: «para hacer una hogaza de pan, mezcla una parte de levadura, dos partes de agua y diez partes de harina...»). Por el contrario, su propósito era mostrarnos cómo los mundos

políticos socialistas, católico-conservadores y liberales habían generado diferentes mundos de bienestar. En sus manos, la RM pretendía ser tan sólo un instrumento contundente usado para remachar suavemente unas enormes diferencias brutas entre grupos de países que se podían haber sacado a la luz empleando tablas y gráficos, sin incurrir en la implicación de los efectos lineales constantes en los países.

¿Cómo podría Esping-Andersen haber sacado el máximo partido a sus datos sin volver a caer en el paradigma estadístico convencional, que entraba en flagrante contradicción con el espíritu de su análisis (y sus críticas explícitas a obras anteriores en esa línea)? Tres estudios anteriores habían sugerido propuestas interesantes. Ragin (1994) llevó a cabo un elaborado análisis ACC empleando siete variables independientes distintas. Se limitó a tomar como variable dependiente la política de pensiones, y realizó análisis separados para los indicadores de cada tipo de régimen. Kangas (1994) comparó los rendimientos del ACC y el análisis de conglomerados con el de las técnicas de regresión tradicionales para contrastar un modelo político simplificado sobre la «calidad» del seguro de enfermedad. Un tercer estudio, realizado por Castles y Mitchell (1992), empleaba simples datos de tablas para construir una tipología alternativa de (cuatro) mundos en el capitalismo de bienestar.

Como podemos observar, los investigadores han estado divididos en dos puntos: cómo definir la variable dependiente, y qué método de análisis emplear. En lo que respecta a la primera de estas cuestiones, (por razones diferentes en cada caso) tanto el estudio de Ragin como el de Kangas se decantan por centrarse en aspectos específicos del régimen de políticas sociales. Sin embargo, una de las tesis más importantes y originales de Esping-Andersen es que los regímenes de políticas sociales son *multidimensionales*: vienen definidos tanto por la magnitud del salario social, como por la calidad de los derechos sociales; incorporan el compromiso del Estado con el empleo, y no sólo las ayudas sociales; y dependen no sólo de las actividades del Estado, sino del equilibrio entre la provisión pública y privada. Castles y Mitchell compartían con Esping-Andersen la ambición de desarrollar un esquema de clasificación comprensivo, pero cuestionaron qué es lo que debería englobar. En su estudio,

el énfasis de Esping-Andersen sobre el aspecto de la desmercantilización de los derechos sociales se sustituye por la consideración de las consecuencias distributivas del salario social ²⁵. Castles y Mitchell dieron el paso decisivo de añadir los impuestos al «mix» de políticas del régimen, pero su análisis prescindió de la distinción entre la provisión pública y la privada, y del compromiso del Estado con el empleo.

Los tres estudios también divergían desde el punto de vista metodológico, comenzando por cómo medían la variable dependiente. Castles y Mitchell siguieron el ejemplo de Esping-Andersen, y presentaron datos de tablas simplificados para apoyar su tipología de regímenes de políticas públicas. Por el contrario, tanto Ragin como Kangas hicieron uso del análisis de conglomerados para asignar los países a los regímenes. Se forzaba a cada país a encajar en un régimen, con lo cual se daba por hecho una cuestión que en realidad requiere una exploración empírica. Surgieron serios problemas prácticos. Kangas tenía dificultades para encontrar los países «liberales», y Ragin quedó en la situación comprometida de tener que asignar un tercio de su países a una categoría «sobrante» que automáticamente los excluía de sus análisis causal (ACC). Ninguno de los investigadores sacó partido de la fuerza potencial que tienen el análisis de conglomerados y otras técnicas multivariantes como mecanismos de reducción de datos. Al esquivar la tesis de Esping-Andersen de que los regímenes de bienestar son multidimensionales, Kangas y Ragin perdieron la oportunidad de contrastar esa aseveración buscando las diferentes dimensiones.

Recordemos que Esping-Andersen identifica cuatro dimensiones diferentes de los Estados de Bienestar: el gasto social, los derechos sociales, la división público / privado y el compromiso con el empleo. Recopiló una multiplicidad de indicadores para las cuatro dimensiones y afirmó que estos indicadores se arraciman en patrones característicos, formando tres constelaciones de políticas públicas distintas. Denominaré a estos regímenes: *institucional* (universalista, desmercantilizador), *residual* (minimalista) y *corporatista* (preservador del estatus). La segunda pata de la teoría de Esping-Andersen postula una unión causal entre cada régimen y el entorno político al que se atribuye su gestación: respectivamente el

socialista, el *liberal* y el *católico-conservador*. Lo que propongo es contrastar estos dos argumentos directamente. ¿Se agrupan los indicadores de las políticas siguiendo las líneas de los tres regímenes típico-ideales? Si es así, ¿hay pruebas de la asociación que postula Esping-Anderson entre el régimen de políticas y el entorno político?

La asociación de los indicadores de políticas públicas puede investigarse a través de diversos métodos. Yo hice pruebas tanto con el análisis de conglomerados, como con el escalamiento multidimensional, para conseguir grupos de *países*. Los resultados fueron bastante consistentes con las expectativas de Esping-Andersen, pero en algunos casos aislados los países iban a parar a grupos de manera extraña e incluso caprichosa. Este es el precio que hay que pagar por intentar encajar los casos en subgrupos mutuamente excluyentes, cuando en realidad la conceptualización de los regímenes de políticas es (y debe ser) típico-ideal. De ahí que yo cambiase de estrategias, intentando vincular indicadores en lugar de casos, e identificando dimensiones subyacentes de las políticas públicas más que grupos de países. Mi razonamiento fue que, analíticamente, la tríada de regímenes de Esping-Andersen se basaba en dos dimensiones de políticas públicas: una representaba la dicotomía de Titmuss (1974) «principios residuales versus principios institucionales»; y la otra comprendía las medidas fragmentadas, jerárquicas y protectoras del estatus, que constituyen anatema tanto entre las fuerzas socialistas como entre las burguesas ²⁶.

Si Esping-Andersen tiene razón y hay tres mundos típico-ideales, deberíamos ser capaces de caracterizar parsimoniosamente los Estados de Bienestar realmente existentes en función de estas dos dimensiones de institucionalismo y corporatismo. En particular, esta caracterización debería aportar información sobre todo en los casos «mixtos», con lo cual ya no estaríamos deseando que desaparecieran de nuestra vista.

Mi reanálisis de 13 de los indicadores de políticas de Esping-Andersen ²⁷ proporciona una confirmación llamativa de su perspectiva sobre los regímenes de bienestar. El escalamiento multidimensional (no reproducido aquí) situó todos los indicadores a lo largo del perímetro de una esfera (una «simplex»), lo cual indica que perte-

necen a un único espectro subyacente. Esto parece reforzar la defensa que hace Esping-Andersen de una conceptualización amplia de las políticas sociales tal y como se había establecido en las hipótesis. La distribución de indicadores específicos siguió las dos dimensiones. Ya que se podía encontrar una solución «limpia» empleando un método que imponía pocas restricciones a la distribución espacial de los objetos, y que, además, no requiere un nivel de medición «alto» (intervalar), había razones para dar un paso más y realizar un análisis factorial. Los factores son combinaciones lineales y económicas de variables. Se generan de tal modo que hay una fuerte correlación entre las variables que exhiben mayores «pesos» en un factor dado, pero una correlación mínima entre factores (lo ideal es que sean «ortogonales») ²⁸.

En la figura 3 se presentan los resultados de un análisis de componentes principales sin rotación. El hecho de que los dos primeros factores den cuenta de la mayoría (casi el 60%) de la varianza es, en sí mismo, una muy buena noticia para la hipótesis de Esping-Andersen. Es evidente que, de derecha a izquierda en la figura, el primer factor capta su dimensión residual/institucional. Exhibe pesos positivos altos en empleo público, gasto en políticas activas de empleo, gasto en igualdad de beneficios y seguridad social; y pesos negativos

altos en asistencia social y en los indicadores de la extensión de la cobertura privada en sanidad y pensiones. El segundo factor (norte-sur) se identifica con la dimensión corporatista de las políticas: pesos positivos altos en el número de esquemas de pensiones y la gran importancia de las pensiones de los funcionarios; y pesos negativos altos en el papel de «las pensiones de los ciudadanos» (Seguridad Social). Aunque los factores no son completamente ortogonales, las zonas de superposición son inteligibles. Por ejemplo, los resultados confirman que los esquemas de pensiones ocupacionales son igualmente repelentes para los regímenes corporatistas y los institucionales. Los resultados también dan a entender que los logros en el ámbito del mercado de trabajo (poco paro y elevada creación de puestos de trabajo) son importantes en el régimen institucional, pero débiles en el corporatista.

Ahora ya estamos preparados para la fase «causal» del análisis, cuyos resultados se presentan en la figura 4. Cuando los 18 países del estudio de Esping-Andersen se ordenan teniendo en cuenta sus puntuaciones en nuestros dos factores, el nexo evidente entre contextos y políticas genera una reveladora cartografía de países. Mientras que Esping-Andersen a veces desdibujaba la distinción entre contextos políticos y regímenes de políticas públicas, en su

Figura 3. Solución bifactorial para los datos de Esping-Andersen

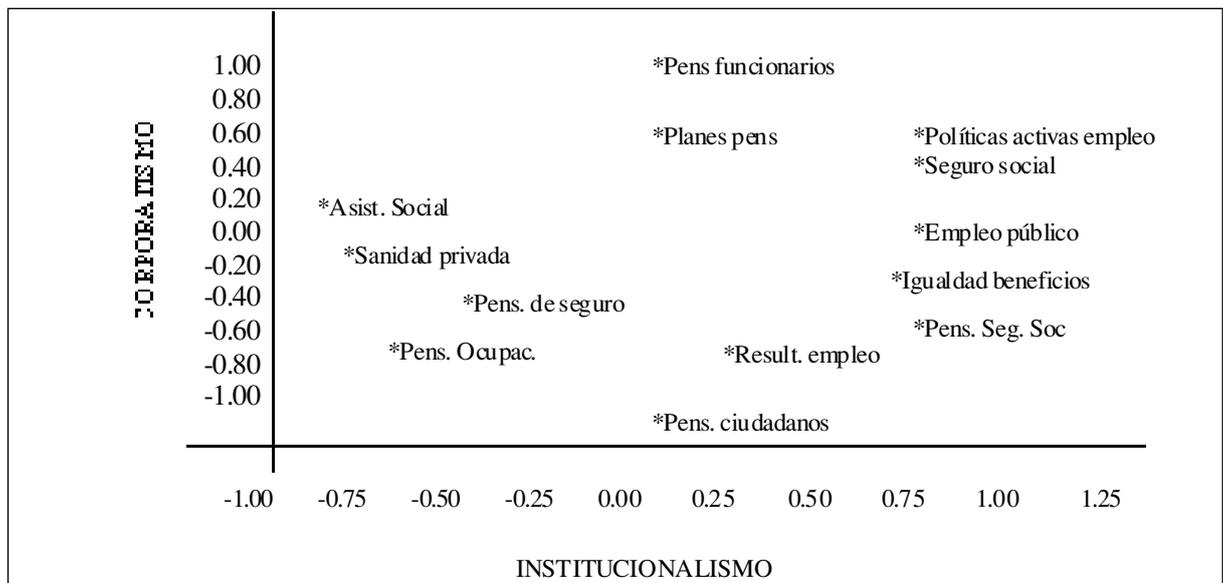
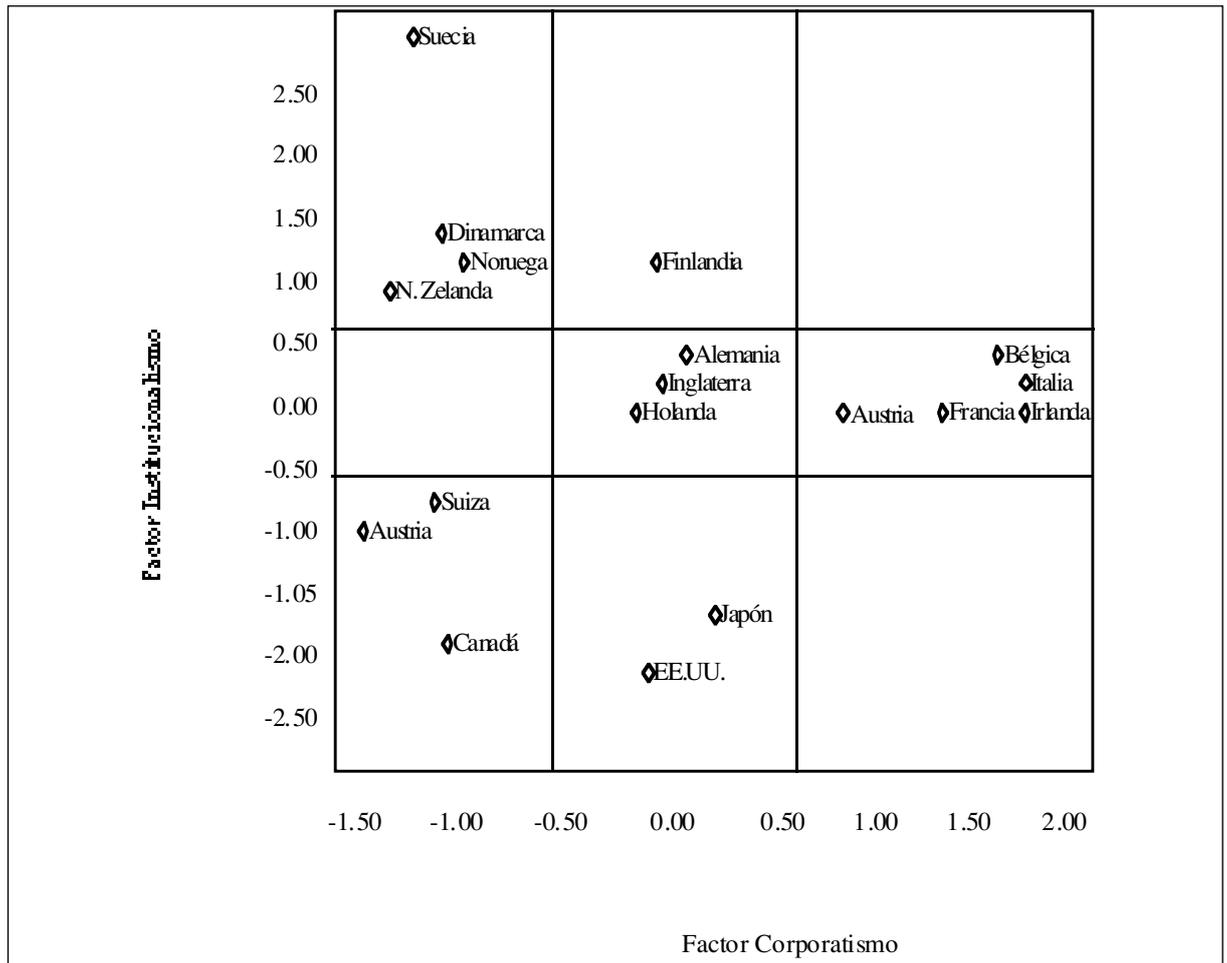


Figura 4. Puntuaciones en los factores de política social para los datos de Esping-Andersen



texto es patente que concibió los tres contextos de la siguiente manera ²⁹:

- *Socialista*: las socialdemocracias escandinavas, caracterizadas por niveles de movilización obrera casi sin parangón en otros países occidentales.
- *Católico-conservador*: naciones europeas –Italia, Francia, Bélgica, Austria e Irlanda– que comparten un pasado absolutista, una democracia relativamente tardía y una población en su mayor parte católica.
- *Liberal*: los EE.UU., Canadá, Suiza y Japón, países en los que la movilización de la clase obrera obrera es muy débil y la herencia conservadora es prácticamente inexistente.

Los cinco países que restan en el estudio de Esping-Andersen son más difíciles de clasifi-

car. Sus tradiciones estatales o bien están próximas al grupo conservador (Alemania y Holanda), o bien son de carácter liberal (Gran Bretaña, Australia y Nueva Zelanda) pero también tienen un pasado de niveles moderados de movilización de la clase obrera.

El ajuste entre los conglomerados políticos y las puntuaciones factoriales es notable, aunque imperfecto. Los Estados liberales y Australia muestran las puntuaciones más negativas en «institucionalismo», mientras que los países escandinavos y Nueva Zelanda obtienen las puntuaciones positivas más altas. Los países restantes, que tienen puntuaciones medias en institucionalismo, se dividen claramente en países conservadores, con puntuaciones altas en el factor corporatismo, y en tres casos mixtos que puntúan casi cero en ambos factores y,

de hecho, quedan fuera del alcance de la teoría. El gráfico también revela una diferencia entre los países no corporatistas: algunos son indiferentes al corporatismo, mientras que otros (bastiones de la socialdemocracia y algunos países liberales) son «repelidos» por aquel. Esto plantea un interesante desafío para el desarrollo ulterior del modelo de Esping-Andersen.

5. Conclusión

A pesar del considerable nivel de debate e innovación metodológicos entre los comparativistas en los últimos años, la RM sigue siendo, con diferencia, el modo predominante de análisis de datos cuantitativos, y la mayor parte de sus críticos ven en el análisis cualitativo (sea formal o convencional) la única alternativa real. Este artículo rechaza por igual ambas tendencias. La aparición y difusión gradual del Análisis Comparativo Cualitativo (ACC) suponen, desde luego, un gran avance para los comparativistas. En su forma más básica, el ACC es un modo de explicitar el tipo de argumentos de «trayectorias causales» característicos de la investigación histórico-comparativa clásica en la línea de Weber, Moore, Rokkan y Skocpol. El deseo de sistematizar la evidencia empírica de ese tipo de argumentos no es nuevo (Somers, 1971). Pero Ragin (1987) es el primero en proponer un procedimiento formal para hacerlo, empleando álgebra booleana para realizar una identificación parsimoniosa de las regularidades subyacentes a una serie de configuraciones de casos.

El ACC no es, desde luego, un método cualitativo, en el sentido de que se base en las habilidades interpretativas de unos analistas que entran en el agua de las descripciones densas mojándose hasta las rodillas. Como han subrayado Griffin y Ragin (1994), si a algo se parece el ACC es a la RM: ambas técnicas aplican reglas que son independientes del investigador, y ambas tratan los casos como «ejemplos múltiples y discretos de fenómenos más generales». En principio el ACC tiene grandes ventajas por su fidelidad a los principios del análisis orientado hacia los casos. Pero, a diferencia de la RM, el ACC está circunscrito a datos categóricos (normalmente

binarios) y utiliza métodos más bien lógicos que estadísticos para generalizar a partir de los casos. Incluso los defensores del ACC admiten que las afirmaciones empíricas que permite deben ser tomadas con cautela. En concreto, como quiera que el método busca causas simples o combinadas que sean *lógicamente* necesarias y/o suficientes, sus resultados son vulnerables tanto al error de medición (incluyendo determinaciones erróneas de presencia/ausencia) como al error de especificación (en el ACC no hay ningún equivalente del término de error). Algunas de estas limitaciones quedaron patentes en mi discusión de la implementación del ACC en la bibliografía sobre regímenes de Estado de Bienestar.

Lo que incorpora el ACC, y lo hace especialmente valioso en el contexto de las macro-comparaciones con *n-pequeños*, es la visibilidad, intimidad y diálogo con los casos —precisamente las ventajas de las que carece la RM. Pero es cuestionable que, para quienes no desean abandonar totalmente la cuantificación, el ACC sea la *única* manera de tener estas ventajas. Mi reanálisis de cuatro estudios basados en la RM sugiere unas cuantas alternativas tanto al ACC como a la RM. Para terminar, incorporo estas sugerencias a una exposición sumaria de las tres opciones más importantes que se les ofrecen a los *cuantis* conscientes de las limitaciones de la RM.

1. *Refinamiento*. Esta es la perspectiva optimista, cuyo mejor ejemplo en esta revisión ha sido la postura de Western respecto al agrupamiento de datos. Western es un buen representante de la estrategia del refinamiento, por el gran potencial que promete el agrupamiento de datos a los comparativistas, y porque es miembro de una nueva generación de economistas políticos cuya destreza estadística es incomparablemente superior a la de los miembros de la generación del autor de este texto. El problema de desagregar de modo inteligible los efectos causales transversales y dinámicos es sólo una de las numerosas cuestiones del análisis de RM para las que los sociólogos y los politólogos han estado buscando inspiración en sus colegas economistas y estadísticos, técnicamente más sofisticados. Pero las limitaciones y dilemas que se descubren a partir de la lectura atenta de la solución de Western al problema de la heterogeneidad causal ofrecen algunos motivos para el

pesimismo. Esta impresión viene respaldada inequívocamente por la revisión de las controversias entre los economistas y sus homólogos en ciencias políticas que ha hecho recientemente Gregory S. Maddala –uno de los decanos más reconocidos de la profesión–. Maddala (1997), muy educadamente –pero sin dejar lugar a dudas–, deja claro que considera que muchas de las modas actuales en la aplicación avanzada de la RM por parte de los politólogos son producto de una imitación equivocada o mal orientada de las manías econométricas.

2. *Triangulación.* John Stephens y sus colegas han ido más allá en la racionalización e implementación del uso de múltiples técnicas de análisis. Sostienen que tanto las historias comparativas no formales de números limitados de casos, como las comparaciones cuantitativas de un espectro amplio de casos que emplean la RM, tienen costes y beneficios que son esencialmente complementarios. En la práctica, adoptan una clara jerarquía epistemológica en la que los resultados de la RM son contrastados con la teoría y el conocimiento de los casos, y las anomalías causales resultantes se someten al seguimiento de trayectorias históricas. Pocos investigadores, y en especial la amplia mayoría de quienes trabajan solos o en escasa colaboración con sus colegas, tienden a seguir este ambicioso *modus operandi*. Es muy típico que la triangulación acabe convirtiéndose en el uso de varias técnicas diferentes para analizar los mismos datos o el mismo problema, y en la constatación de si generan resultados parecidos. Tanto el estudio de Kangas, citado en la sección 4, como el trabajo de Ebbinghaus y Visser (en prensa) sobre tendencias comparativas en la afiliación sindical, han combinado el ACC con la RM. Al igual que la combinación que hacen Hall y Franzese del análisis de casos y de tablas con la RM, este tipo de análisis híbrido es una innovación que se agradece, especialmente cuando es resultado de un diálogo auténtico entre casos y generalizaciones. Así como este diálogo lo facilitan tanto el ACC como los métodos de tablas, las políticas restrictivas de los editores de revistas, y otros factores, pueden limitar fácilmente este potencial en la práctica.

3. *Sustitutos.* La tercera estrategia, y la más radical, para hacer frente a los problemas de la

RM sin abandonar el análisis cuantitativo es entregarse por completo a técnicas alternativas. Las secciones 2 y 3 de este artículo presentaron algunas tablas de contingencia y diagramas de flujo sencillos en los que se identifica a los países y los datos numéricos se agrupan en nuevas categorías o son promediados. Se mostró que, de hecho, estas técnicas pueden superar algunas de las limitaciones menos atractivas de la RM, incorporando al mismo tiempo elementos clave del estudio de casos. Por un lado, daban cabida a ideas analíticas complejas (esto es, síndromes y jerarquías de causas). Por otro lado, llamaban nuestra atención sobre casos que merecían un nuevo escrutinio comparativo más detallado –algo de lo que ningún otro método puede hacer gala. También he sugerido que, dado que se adaptan mejor a nuestros supuestos ontológicos, no hay ninguna razón por la cual los métodos estadísticos multivariantes distintos de la regresión no deban ser explotados por los comparativistas. El ejemplo que se ha puesto aquí es el de la utilidad del análisis factorial para poner en claro la evidencia empírica del enfoque de Esping-Andersen sobre la diversidad de Estados de Bienestar ; pero el análisis de conglomerados, el escalamiento multidimensional y los métodos de reducción de datos desarrollados por Dirk Berg-Schlesser y sus colaboradores son otros mecanismos mediante los cuales el «handicap» tradicional del *n pequeño* y las muchas *varindep*s (variables independientes) puede convertirse en un activo importante.

NOTAS

¹ Artículo redactado para su presentación en la Segunda Conferencia sobre «The Welfare State at the Crossroads (El Estado de Bienestar en la encrucijada)», Estocolmo, 12-14 de junio de 1998. Debo al entorno agradable y la buena voluntad recibidas de Walter Korpi, así como a la alentadora correspondencia con Bruce Western, el ánimo para poner por escrito estas ideas. Frank Castles, Peter Hall, Robert Franzese y los participantes en el «Wokshop on Economic Internationalization» en la Universidad de Duke (25 al 27 de mayo de 1997) y en el seminario del Wissenschaftszentrum de Berlín en febrero de 1999 realizaron comentarios a borradores previos que me fueron de gran ayuda.

² El *American Journal of Political Science* se ha constituido en un foro puntero para los trabajos más vanguardistas sobre cómo emplear la RM de manera más sofisticada.

³ Desde luego, es discutible en qué medida el universo de investigación es restringido. Los criterios del tipo

«países capitalistas avanzados con politeyas democráticas estables y tamaños de población no triviales» pueden ser arbitrarios desde el punto de vista teórico, y deberían ser siempre susceptibles de discusión. Hay una antigua (aunque de ninguna manera universal) tendencia a incluir a Grecia, España y Portugal tras su democratización (y desde el punto de vista más práctico, tras su incorporación en las bases de datos de la OCDE). Otros candidatos a entrar en los trabajos de economía política comparativa sobre lo hasta ahora conocido como «países occidentales» podrían estar en los antiguos estados del bloque soviético, Latinoamérica y Asia Oriental. Hay buenos argumentos tanto a favor como en contra de la expansión del universo de los estudios comparativos. Por ejemplo, compárese Geddes (1990) y Boyer (1997).

⁴ Incluso la conocida recomendación de Przeworski/Teune (1970) de que los comparativistas deberían esforzarse en sustituir los nombres propios de los países por los nombres abstractos de las variables no descartaba completamente esta visión. Vale la pena recordar que Przeworski y Teune estaban protestando amargamente contra el predominio de los especialistas en «estudios de área» en la ciencia política comparativa, e instaban a sus colegas a evitar argumentos de tipo particularizador que podían ser una camisa de fuerza tanto para la teoría como para la comparación. Muchos defensores contemporáneos del análisis de casos, incluido Ragin, no habrían tenido nada en contra de esta postura.

⁵ La crítica en este punto no es que la cuantificación sobresimplifique la realidad compleja. Tanto si en el análisis se usan medidas cuantitativas o representaciones narrativas, en la investigación social siempre hay un dilema entre la exactitud y la parsimonia. La cuestión es que el empleo de la RM fomenta lo que bien podría ser una creencia equivocada en que nuestras medidas son precisas y continuas.

⁶ Sorprendentemente, pocos practicantes de la RM tienen conocimiento de la polémica sobre si los controles estadísticos realmente son capaces cumplir su función. Estoy en deuda don Bruce Western por dirigirme hacia esta bibliografía.

⁷ Recientemente, Abbot (1997) ha propuesto una elegante formulación de este problema. Los análisis de variables «intentan entender el proceso social desarrollando transformaciones lineales de un espacio hiperdimensional (de “efectos principales” y, de vez en cuando, de interacciones entre ellos) en una única dimensión (la variable dependiente)... Ahora bien, esta estrategia (...) es útil sólo si el espacio de los datos está lleno más o menos uniformemente» (Abbot, 1997).

⁸ Agradezco a Rob Franzese el haberme forzado a darle vueltas a este punto.

⁹ Un ejemplo es el argumento de Theda Skocpol de que son «las capacidades estatales» y «los legados de las políticas», y no los recursos de la clase trabajadora, lo que explica las variaciones políticas entre EE.UU. y otros países (p.ej., Weir y Skocpol, 1985). Sin embargo se ha demostrado que la teoría centrada en el Estado no se puede contrastar (o sólo muy superficialmente) mediante la investigación comparativa cuantitativa. En parte, como consecuencia de lo resbaladizo de la teoría, y, en parte, debido a la ausencia de buenos indicadores, la mayor parte de sus variables independientes no se pue-

den ni cuantificar ni simplificar. Podemos medir aspectos sencillos de la estructura de los Estados, como p.ej., las provisiones constitucionales, incorporadas en la obra de Huber, Ragin y Stephens (1993) sobre las variaciones en el Estado de Bienestar. Excepto en la investigación histórica no formal, el enfoque de la acción y la agenda políticas, las capacidades estatales, los legados políticos y las iniciativas autónomas de los «managers» del Estado no ha sido tomado muy en serio. Una excepción es la regresión y el estudio ACC de los estados norte americanos realizado por Amenta y Poulson (1996). Sin embargo, es la excepción que confirma la regla, ya que la medición de conceptos como «la fuerza administrativa» sólo fue posible en este estudio porque comparaba unidades sub-nacionales de una entidad nacional uniforme.

¹⁰ Rothstein reconoce sabiamente una limitación más fundamental de la estática comparativa cuando admite que para demostrar su hipótesis debe ir más allá del análisis de corte transversal y aportar pruebas históricas concretas. Aquí me preocupa sólo el empleo que hace Rothstein de la RM para presentar pruebas de corte transversal a favor de su argumento.

¹¹ No parece lógico que Rothstein incluyese Islandia (con sólo 80.000 afiliados potenciales), y no Luxemburgo, Mónaco o Andorra en su análisis. El problema del indicador de representación de los partidos de izquierda de Rothstein, que éste tomó prestado de Wilensky (1981) es que cubre todo el período 1919-1979. Las interrupciones y discontinuidades debidas a los períodos no democráticos y bélicos vician enormemente los datos anteriores a 1945. En el otro extremo del espectro, los datos de afiliación revelan que las diferencias internacionales se estabilizaron después de 1965. Por lo tanto, parece razonable considerar las dos primeras décadas posteriores a la guerra como el período de formación de las políticas. Las cifras de la fuerza media de la izquierda en el gobierno en este período se tomaron de la serie de datos empleados por Korpi y Shalev (1980). Ninguna de estas modificaciones desvirtúan los resultados que el propio Rothstein obtuvo utilizando la RM. Por el contrario, empleando los datos modificados, el valor del coeficiente beta para la variable Gante aumenta a expensas de la de la fuerza de la izquierda.

¹² Sencillamente, se dicotomizó la afiliación sindical después de que las tablas exploratorias revelasen que tenía un evidente efecto umbral sobre la afiliación. Con la excepción de Suiza y Holanda, el resto de los países pequeños (no más de 5 millones de afiliados potenciales) tenía una densidad de más del 50 por ciento; mientras que todos los países grandes (más de 10 millones) puntuaban menos del 50 por ciento. Dentro de estas dos categorías no había ninguna relación perceptible entre las dos variables.

La fuerza de la izquierda se agrupó en cuatro categorías que reflejan cortes en su distribución. «Ninguna», para los casos con representación de partidos de la izquierda cero o despreciable (hasta el 4% de Japón); «débil», de 7 a 15%; «media», de 22 a 29%, más un caso intermedio (Gran Bretaña) con un 36%; y «fuerte», 45% o más.

¹³ Por otro lado, la fuerza de la izquierda sólo diferenció débilmente los niveles de afiliación de los países pequeños, y no los de los países grandes (excepto quizás en el caso británico).

¹⁴ La importancia que tienen este tipo de anomalías en el progreso científico lo ha ilustrado recientemente Rogowski (1995). Se observará que estos datos también indican anomalías que pueden no tener relevancia para el interés de Rothstein por el sistema Gante, pero que deberían interesar a los estudiosos de la afiliación sindical. Una es la gran distancia entre EE.UU. y Canadá; otra es la densidad sindical irlandesa, inesperadamente alta.

¹⁵ Visser (1992) ha señalado que la mayor parte de las diferencias entre el sistema sindical belga y holandés pueden deberse al hecho de que los sindicatos holandeses no tienen presencia en el lugar de trabajo. Los orígenes del hiato entre Noruega y Suecia son menos claros, pero pueden retrotraerse a la menor eficacia del movimiento sindical noruego en algunos de los sectores que crecieron a partir de los años sesenta, cuando la densidad noruega se estancó, mientras que la sueca comenzó un largo período de crecimiento. Los datos recogidos por D'Agostino revelan distancias considerables en la densidad sindical a favor de Suecia en las siguientes categorías (en parte interrelacionadas): mujeres, comercio y servicios de titularidad privada, y trabajadores de cuello blanco.

¹⁶ Todas las economías coordinadas (.75 ó 1) excepto Dinamarca tenían tasas de desempleo inferiores al 3.5%, mientras que en todas las no coordinadas (0 ó .25), excepto Nueva Zelanda, superaban el 3.5%.

¹⁷ Goldthorpe recomienda aún más encarecidamente que los investigadores amplíen «el abanico geográfico y sociocultural» de su investigación. Sin embargo, no se puede decir —como tampoco, del agrupamiento de datos— que la solución recomendada en esta cuestión sea muy popular. No se trata tan sólo (como admite Goldthorpe) de que la calidad y disponibilidad de los datos sea limitada fuera del bloque —países de la OCDE— que interesa a la audiencia a la que él —y yo— nos dirigimos. Más bien, es de sobra conocido que, en general, los llamados «costes de especificación» en que se incurre por ir más allá de la OCDE (factores causales añadidos y senderos causales alternativos) son mayores los beneficios potenciales. En cualquier caso, dista de estar claro que, en la práctica, aumentar el tamaño de la muestra sirva para superar las dificultades que entraña la comparación internacional como ciencia exacta. Pensemos en el ejemplo de la economía del crecimiento. Incluso en un campo como este, desarrollado desde el punto de vista teórico, y en el que fue posible reunir datos comparables para el pasmoso número de 119 países, Levine y Renelt (1992) se vieron incapaces de escoger entre varias teorías empleando la RM.

¹⁸ Para una exposición anterior y no técnica, véase Stimson (1985). Para elaboraciones más complejas, así como para referencias a numerosas aplicaciones del agrupamiento de datos en las áreas de la política social y la economía política comparativas, véase Hicks (1994) y Beck y Katz (1995).

¹⁹ Recientemente, Beck y Katz (1995) descalificaron muchos de los estudios punteros en este género, algunos de ellos realizados por famosos investigadores con prodigiosas habilidades polimétricas. La conclusión de Beck y Katz (que el mejor enfoque en muchos casos es la regresión MCO estándar, con la variable dependiente retardada incluida en el lado derecho) contradice el saber convencional. Véase, por ejemplo, Sayers (1989).

²⁰ En su propio intento de hacer esto, Hicks (1994) utilizó periodicidades (datos anuales versus medias quinquenales) y métodos (Parks/GLS versus «componentes de varianza») diferentes para lo que él llamaba modelos «a corto plazo» y «a medio plazo». Sin embargo, si he entendido su artículo correctamente (y bien pudiera ser que no), en realidad Hicks no comparó resultados de corte transversal y de series temporales. También parece que no tuvo en cuenta la posible inestabilidad de los parámetros internacionales o temporales.

²¹ Los investigadores a menudo contemplan el uso de variables ficticias para captar efectos de país o de período. Sin embargo, lo que las variables ficticias captan realmente son las diferencias en la constante o «valor base» de la variable dependiente. (Como consecuencia, son culpables de «robar» los efectos de variables institucionales como el corporatismo.) Para comprobar si hay diferencias de países o períodos en las pendientes, se necesitarían términos de interacción, mucho más costosos en términos de grados de libertad. Un término medio, más sensible al contexto, pero menos exhaustivo en términos de grados de libertad, sería permitir que tanto las constantes como los parámetros variasen en contextos generalizados. Un innovador estudio de O'Connell (1994) ha seguido esta estrategia, incluyendo términos de interacción para países «corporatistas» versus «no corporatistas», y para períodos «de crisis» versus «sin crisis».

²² Western también propone emplear distribuciones «de cola gruesa» para mejorar la precisión, incluyendo casos atípicos pero minimizando sus efectos. Aquí sólo estoy interesado en su empleo de la modelización jerárquica para estimar efectos contextuales sobre parámetros de series temporales.

²³ En concreto, tres de los cuatro países escandinavos tienen coeficientes positivos. En el otro (Finlandia), es cero.

²⁴ Western (1996) admite que la variable de gobierno de la izquierda en Japón era constante, y previene contra una «interpretación sustantiva» del resultado japonés. El problema es que, independientemente de lo que él o yo podamos hacer con este resultado, el modelo jerárquico bayesiano se basa necesariamente en él para establecer influencias contextuales en la dinámica socialdemócrata.

²⁵ Castles y Mitchell señalan que las medidas de desmercantilización y universalismo de Esping-Andersen habrían «penalizado» a Australia y Nueva Zelanda por tener políticas basadas en la comprobación de recursos. Su contra-argumento era que, en la práctica, esta comprobación excluye sólo a los pudientes, y que, en combinación con impuestos progresivos para financiar los beneficios, tiene un efecto igualitario que no podían aprehender las mediciones de Esping-Andersen para los casos de las Antípodas. Haciéndome eco de estas críticas al tratamiento de Esping-Andersen de la comprobación de recursos, en mi propio análisis me abstengo de utilizar sus índices de desmercantilización o su escala de universalismo para los programas de desempleo y enfermedad.

²⁶ La innovación de *Los tres mundos* consiste precisamente en que reconoció que las variaciones en el Estado de Bienestar no podían describirse simplemente recurriendo a la polaridad entre las políticas «institucionales» pro-obreras y el Estado de Bienestar «residual», favorecido por el liberalismo económico de la

derecha. La triple comparación entre Austria, Alemania y Suecia reveló que en Austria el impacto de la fuerza del movimiento obrero era contrarrestado por una derecha fuerte, la influencia católica y una historia de intentos por parte de los regímenes conservadores de emplear la política social para dividir a la clase obrera (Esping-Andersen, 1984). A esto se debe que el Estado de Bienestar austríaco se asemeje más al alemán que a la versión sueca. Esta intuición vino anticipada de referencias de obras anteriores a la influencia del catolicismo y de la unidad de la derecha política (Castles, 1978; Korpi, 1978; Stephens, 1979). Lo que Esping-Andersen hizo fue cristalizar la dimensión de la política «corporatista» situándola concretamente en un entorno europeo particular.

²⁷ Los datos se obtuvieron de la manera siguiente. Referidas a Esping-Andersen (1990): salario social (tabla 5.1, fuente de datos del autor); universalismo (tabla 3.1, datos desagregados mediante un programa del autor); desmercantilización (tabla 2.1.); uniformidad (tabla 3.1); asistencia social (tabla 3.1); división público-privado en la sanidad (tabla 3.1) y las pensiones (tabla 4.3); «realización del pleno empleo» (tabla 5.9, datos del autor). Los gastos del programa de políticas activas de empleo sobre el PIB (c. 1975) y de empleo público en porcentaje sobre el total del empleo (1980) se mencionan en Esping-Andersen (1990) y se analizan en Esping-Andersen (1985), pero los datos los obtuvo el autor directamente.

²⁸ Así, el investigador espera que cada ítem pese mucho en sólo uno de los factores. El procedimiento conocido como «rotación» de factores está diseñado para hacer que esto ocurra, pero yo opté por la contrastación más difícil de un análisis sin rotación.

²⁹ Se observará que, con una excepción (la asignación de Finlandia al grupo socialdemócrata), la clasificación que deduje del análisis discursivo de Esping-Andersen puede reproducirse perfectamente mediante la agrupación de países en función de sus tres indicadores políticos: movilización de la clase obrera, absolutismo / democratización tardía, y catolicismo.

BIBLIOGRAFÍA

- ABBOTT, Andrew (1988): «Transcending General Linear Reality», *Sociological Theory*, 6 (2): 169-86.
- ABBOTT, Andrew (1997): «On the Concept of Turning Point», *Comparative Social Research*, 16:85-105.
- AMENTA, Edwin and Jane D. POULSEN (1996): «Social Politics in Context: The Institutional Politics Theory and Social Spending at the End of the New Deal», *Social Forces*, 75 (1): 33-59.
- BECK, N. and J. N. KATZ (1995): «What to do (and not to do) with time-series cross-section data», *American Political Science Review*, 89 (3): 634-47.
- BERG-SCHLOSSER, Dirk and Giselle DE MEUR (1994): «Conditions of Democracy in Interwar Europe- a Boolean Test of Major Hypotheses», *Comparative Politics*, 26 (3): 253-79.
- BOYER, Robert (1997): «Comparing Germany and Japan: Methodological Issues, Main Findings from the Regulation Approach, and an Agenda for Further Research», texto preparado para la conferencia «Germany and Japan: The Future of Nationally Embedded Capitalism in a Global Economy», University of Washington, Seattle, April 10-12.
- CASTLES, Francis G. (1978): *The Social Democratic Image of Society*, Londres, Routledge.
- CASTLES, Francis Geoffrey (1993): *Families of Nations: Patterns of Public Policy in Western Democracies*, Brookfield VT, Dartmouth.
- CASTLES, Francis G. and Deborah MITCHELL (1992): «Identifying Welfare State Regimes: The Links between Politics, Instruments, and Outcomes», *Governance*, 5 (1): 1-26.
- EBBINHAUS, Bernard and Jelle VISER (1996): «Social Change and Unionization: The Development of Union Membership in Western Europe, 1950-90», Annual Meeting of the American Political Science Association, San Francisco, August 29-September 1.
- ESPING-ANDERSEN, Gosta (1985): «Power and Distributional Regimes», *Politics and Society*, 14 (2): 223-56.
- ESPING-ANDERSEN, Gosta (1990): *The Three Worlds of Welfare Capitalism*, Cambridge, Polity Press.
- ESPING-ANDERSEN, Gosta and Walter KORPI (1984): «Social Policy as Class Politics in Post-War Capitalism: Scandinavia, Austria, and Germany», pp. 179-208, en J. Goldthorpe (Ed.): *Order and Conflict in Contemporary Capitalism*, Oxford University Press.
- FIREBAUGH, Glenn (1980): «Cross-National Versus Historical Regression Models: Conditions of Equivalence in Comparative Research», *Comparative Social Research*, 3: 333-44, Paper presented to the 73rd Annual Meeting of the American Sociological Association, San Francisco, September 4-8.
- FRANZESE, Robert J. (en prensa): «The Political Economy of Over-Commitment: A Comparative Study of Democratic Management of the Keynesian Welfare State», tesis doctoral, Political Science, Harvard University, Cambridge MA.
- FREEDMAN, David A. (1991): «Statistical Models and Shoe Leather», *Sociological Methodology*, 21: 291-313.
- GEDDES, Barbara (1990): «How the Cases you Choose Affect the Answers You Get: Selection Bias in Comparative Politics», *Political Analysis*, 2: 131-50.
- GOLDTHORPE, J. H. (1991): «The Uses of History in Sociology-Reflections On Some Recent Tendencies», *British Journal of Sociology*, 42 (2): 211-30.
- GOLDTHORPE, John H. (1997): «Current Issues in Comparative Macrosociology: A Debate on Methodological Issues», *Comparative Social Research*, 16: 1-26.
- GRIFFIN, L. and C. C. RAGIN (1994): «Some Observations On Formal Methods of Qualitative Analysis», *Social Methods & Research*, 23 (1): 4-21.
- GRIFFIN, Larry and Pamela BARNHOUSE WALTERS and Philip O'CONNELL and Edward MOOR (1986): «Methodological Innovations in the Analysis of Welfare-State Development: Pooling Cross Sections and Time Series», pp. 101-38 en *The Futures for the Welfare State*, coordinado por N. Furniss, Bloomington, Indiana University Press.
- GRIFFIN, Larry and Philip J. O'CONNELL/Holly J. MCCAMMON (1989): «National Variation in The Context of Struggle: Post-war Class Conflict and Market

- Distribution in the Capitalist Democracies», *Canadian Review of Sociology and Anthropology*, 26: 37-68.
- HALL, Peter A./Robert J. Jr. Franzese (1998): «Mixed Signals: Central Bank Independence, Coordinated Wage Bargaining, and European Monetary Union», *International Organization* (verano), en prensa.
- HIBBS, Douglas A., Jr. (1976): «Industrial Conflict in Advanced Industrial Societies», *American Political Science Review*, 70 (4): 1033-58.
- HICKS, Alexander (1994a): «Introduction to Pooling», pp 169-88, en *The Comparative Political Economy of the Welfare State*, coordinado por T. Janoski/A. M. Hicks, Cambridge, Cambridge University Press.
- HICKS, Alexander (1994b): «The Social Democratic Corporatist Model of Economic Performance in the Short- and Medium-Run Perspective», pp 189-217, en *The Comparative Political Economy of the Welfare State*, coordinado por T. Janoski/A. M. Hicks. Cambridge, Cambridge University Press.
- HUBER, Evelyn/Charles RAGIN/John D. STEPHENS (1991): «Quantitative Studies of Variation Among Welfare States: Towards a Resolution of the Controversy», ISA Conference on Comparative Studies of Welfare State Development, Helsinki.
- HUBER, E./C. RAGIN/J. D. STEPHENS (1993): «Social Democracy, Christian Democracy, Constitutional Structure, and the Welfare State», *American Journal of Sociology*, 99 (3): 711-49.
- IVERSEN, Torben (1998): «Wage Bargaining, Hard Money and Economic Performance: Theory and Evidence for Organized Market Economies», *British Journal of Political Science*, 28: 31-61.
- JANOSKI, Thomas/Alexander M. HICKS (coord.) (1994): *The Comparative Political Economy of the Welfare State*, Cambridge, Cambridge University Press.
- JANOSKI, Thomas/Christa MCGILL/Vanessa TINSLEY (1997): «Making Institutions Dynamic in Cross-National Research: Time-Space Distancing in Explaining Unemployment», *Comparative Social Research*, 16: 227-68.
- KANGAS, Olli (1994): «The Politics of Social Security: On Regressions, Qualitative Comparisons and Cluster Analysis», pp 346-64, en *The Comparative Political Economy of the Welfare State*, coordinado por T. Janoski/A. M. Hicks, Cambridge, Cambridge University Press.
- KING, Gary/Robert O. KEOHANE/Sidney VERBA (1994): *Designing Social Inquiry: Scientific Inference in Qualitative Research*, Princeton NJ, Princeton University Press.
- KISER, E./M. HECHTER (1991): «The Role of General Theory in Comparative-Historical Sociology», *American Journal of Sociology*, 97 (1): 1-30.
- KORPI, Walter (1978): *The Working Class in Welfare Capitalism: Work, Unions and Politics in Sweden*, London, Routledge and Kegan Paul.
- KORPI, Walter/Michael SHALEV (1980): «Strikes, Power and Politics in Western Nations, 1900-1976», *Political Power and Social Theory*, 1: 301-34.
- LEAMER, Edward E. (1983): «Let's Take the Con out of Econometrics», *American Economic Review*, 73 (1): 31-44.
- LIEBERSON, Stanley (1985): *Making It Count*, Berkeley CA, University of California Press.
- LIEBERSON, S. (1991): «Small N's and Big Conclusions: An Examination of the Reasoning in Comparative Studies Based On a Small Number of Cases», *Social Forces*, 70 (2): 307-20.
- LIEBERSON, S. (1994): «More On the Uneasy Case For Using Mill-Type Methods in Small-N Comparative Studies», *Social Forces*, 72 (4): 1225-37.
- LUSTICK, Ian (1996): «History, Histograms, and Political Science: Multiple Historical Records and the Problem of Selection Bias», *American Political Science Review*, 90 (3): 605-18.
- MADDALA, G. S. (1997): «Recent Developments in Econometric Modelling: A Personal Viewpoint», Annual Meeting of the Political Methodology Group, Columbus OH, July 24-27, http://wizard.ucr.edu/polmeth/working_papers97/madda97.html.
- MARTIN, Andrew (1973): *The Politics of Economic Policy in the United States: A Tentative View from a Comparative Perspective*, Beverly Hills, Sage, (Sage Professional Papers in Comparative Politics, No. 01-040).
- MJØSET, Lars/et al. (eds.) (1997): *Methodological Issues in Comparative Social Science*, Greenwich CT, JAI Press.
- O'CONNELL, Philip (1994): «National Variation in the Fortunes of Labor: A Pooled and Cross-Sectional Analysis of the Impact of Economic Crisis in the Advanced Capitalist Nations», pp. 218-42, en *The Comparative Political Economy of the Welfare State*, coordinado por T. Janoski/M. Hicks, Cambridge: Cambridge University Press.
- PALDAM, Martin/Peder J. PEDERSEN (1982): «The Macroeconomic Strike Model: A Study of 17 Countries, 1948-1975», *Industrial and Labor Relations Review*, 35 (4): 504-21.
- PRZEWORKSI, Adam/Henry TEUNE (1970): *The Logic of Comparative Social Inquiry*, New York, Wiley.
- RAGIN, Charles (1994): «A qualitative comparative analysis of pension systems», pp. 320-45, en *The Comparative Political Economy of the Welfare State*, coordinado por T. Janoski/M. Hicks, Cambridge, Cambridge University Press.
- ROGOWSKI, Roland (1995): «The Role of Theory and Anomaly in Social-Scientific Inference», *American Political Science Review*, 89 (2): 467-70.
- ROTHSTEIN, Bo (1990): «Labour Market Institutions and Working Class Strength», pp. 33-56, en Steinmo, K. Thelen/F. Longstreth (coords.): *Structuring Politics: Historical Institutionalism in Comparative Analysis*, Cambridge, University Press.
- SAYRS, Lois (1989): *Pooled Time Series Analysis*, London, Sage (Quantitative Applications in the Social Sciences, No. 70).
- SHALEV, Michael (1978): «Lies, Damned Lies and Strike Statistics: The Measurement of Trends in Industrial Conflict», pp. 1-19, 322-34 en *The Resurgence of Class Conflict in Western Europe since 1968*, coordinado por C. Crouch/A. Pizzorno, London, Macmillan.
- SHALEV, Michael (1979): «The Strike: Theoretical and Empirical Studies of Industrial Conflict Across Time and Nations», manuscrito no publicado, Madison, University of Wisconsin.

- SHALEV, Michael (1980): «Trade Unionism and Economic Analysis: The Case of Industrial Conflict», *Journal of Labor Research*, 1 (1): 133-73.
- SHALEV, Michael (1983): «Class Politics and the Western Welfare State», pp. 27-50, en *Social Policy Evaluation: Social and Political Perspectives*, coordinado por S. E. Spiro and E. Yuchtman-Yaar, New York, Academic Press.
- SHALEV, Michael (1990): «Class Conflict, Corporatism and Comparison: The Japanese Enigma», pp. 60-93 en *Japanese Models of Conflict Resolution*, coordinado por S. N. Eisenstadt/E. Ben-Ari, Londres, Kegan Paul International.
- SNYDER, David (1975): «Institutional Setting and Industrial Conflict: Comparative Analyses of France, Italy and the United States», *American Sociological Review*, 40 (3): 259-78.
- SOMERS, Robert H. (1971): «Applications of Expanded Survey Research Model to Comparative Institutional Studies», pp. 357-420 en *Comparative Methods in Sociology: Essays on Trends and Applications*, coordinado por I. Vallier. Berkeley (CA), University of California Press.
- STEPHENS, John D. (1979): *The Transition from Capitalism to Socialism*, London, Macmillan.
- STIMSON, James A. (1985): «Regression in Space and Time: A Statistical Essay», *American Journal of Political Science*, 29: 915-47.
- TITMUS, Richard Morris (1974): *Social Policy*, London, George Allen and Unwin.
- VISSER, Jelle (1992): «The Strength of Union Movements in Advanced Capitalist Democracies: Social and Organizational Variations», en *Labour Movements towards the Year 2000*, coordinado por M. Regini, London, Sage.
- WEIR, Margaret/Theoda SKOCPOL (1985): «State Structures and the Possibilities for “Keynesian” Responses to the Great Depression in Sweden, Britain and the United States», pp. 107-63, en *Bringing The State Back In*, coordinado por P. B. Evans/D. Rueschemeyer/T. Skocpol, Cambridge, Cambridge University Press.
- WESTERN, B. (1995): «Concepts and Suggestions For Robust Regression Analysis», *American Journal of Political Science*, 39 (3): 786-817.
- WESTERN, Bruce (1996): «Casual Heterogeneity in Comparative Research: A Bayesian Hierarchical Modelling Approach» trabajo no publicado, Department of Sociology, Princeton University.
- WESTERN, B./S. JACKMAN (1994): «Bayesian Inference For Comparative Research», *American Political Science Review*, 88 (2): 412-23.
- WILENSKY, Harold L. (1981): «Leftism, Catholicism, and Democratic Corporatism: The Role of Political Parties in Recent Welfare State Development», pp. 345-82 en *The Development of Welfare States in Europe and America*, coordinado por P. Flora/A. J. Heidenheimer, New Brunswick NJ, Transaction.