

EL MODELO ECONÓMICO COMO CIENCIA BASURA

Ted Goertzel

Universidad de Rutgers

RESUMEN

Es muy habitual encontrar investigaciones publicadas que con los mismos datos llegan a resultados muy diferentes. Las técnicas de regresión múltiple y otros modelos matemáticos no han cumplido su promesa de ser un método de evaluación de las políticas públicas. En este artículo, se definen como ciencia basura los modelos de regresión que no han demostrado que funcionen con datos reales, distintos a los datos utilizados para originarlos. Muchos científicos sociales han dedicado años y años a aprender y enseñar los modelos de regresión. Aún continúan utilizándolos para establecer relaciones causales que no están justificadas por sus datos, pero que se repiten con frecuencia en los argumentos políticos. A estos argumentos los llamé los mitos de la regresión múltiple.

ABSTRACT

It is quite common rival studies to be published using econometric methods to reach opposite conclusions about the same set of data. Multiple regression and other mathematical modeling techniques have simply not lived up to their promise as a method of evaluating the impact of public policies. Regression models that have not been demonstrated to work with fresh data, other than the data used to create them, are junk science. Many social scientists have devoted years to learning and teaching regression modeling. They continue to use regression to make causal arguments that are not justified by their data, but that get repeated over and over in policy arguments. I call these arguments the myths of multiple regression.

Key words: social sciences, impact of public policies, junk science, myths of multiple regression

¿Cree que por cada preso que es ejecutado en los Estados Unidos, se evitan ocho nuevos asesinatos? ¿Cree que el aumento del 1% de ciudadanos que dispone de un arma reduce en un 3.3% la tasa de asesinatos de un estado? ¿Cree que entre el 10 y 20% del descenso de la criminalidad en la década de 1990 se debe al incremento de los abortos en los años setenta? ¿O que la tasa de asesinatos podría haber aumentado en un 250% desde 1974, si los Estados Unidos no hubiera construido nuevas prisiones? ¿Cree en la predicción de que la reforma de bienestar de los años noventa llevará a 1.100,000 niños a la pobreza?

Si usted se introduce en cualquiera de estos estudios, puede haber caído en un tipo nefasto de ciencia de basura: el uso de modelos matemáticos para evaluar el impacto de políticas sociales. Estos estudios son aparentemente muy llamativos. Son realizados por reconocidos científicos sociales de instituciones prestigiosas y suelen publicarse en las revistas científicas. Están llenos de cálculos estadísticos demasiados complejos para cualquiera otro especialista que quiera resolverlos. Proporcionan *hechos* numéricos precisos que se citan a menudo en los debates de política. Pero estos *hechos* no son más que voluntad, simples deseos. Con mucha frecuencia antes de que se haya cerrado definitivamente el estudio, aparece otro con *hechos* igualmente precisos e impresionantes, pero completamente diferentes. A pesar de su precisión numérica, estos *hechos* no tienen más validez que las visiones de los adivinos.

Algunas de estas predicciones se basan en la técnica de regresión múltiple, que utiliza el análisis de correlaciones para realizar explicaciones causales. Aunque los economistas son los creyentes principales de este hermético arte, los sociólogos, criminólogos y otros científicos sociales también lo utilizan. Es conocido bajo distintos nombres como “diseños econométricos” “modelos estructurales,” “análisis de una vía” o simplemente “análisis multivariante”. Todos son formas de utilizar los datos de correlaciones para realizar explicaciones causales.

El problema, como sabe cualquiera que haya manejado estadística, es que la correlación no significa causalidad. Una correlación entre dos variables puede estar *adulterada* cuando está causada por una tercera variable. Los investigadores de la regresión múltiple intentan superar el problema de esta distorsión incluyendo simplemente todas las variables en los datos de análisis. Los datos disponibles para este propósito no depende de esta tarea y los estudios han fracasado de forma consistente. Pero muchos científicos sociales han dedicado muchos años a aprender y enseñar el modelo de regresión múltiple. Continúan utilizando la regresión para hacer argumentos causales que no están justificados por sus datos, pero que se repiten una y otra vez en los argumentos políticos. Yo llamo a estos argumentos los mitos de regresión múltiple.

Cinco mitos de la Regresión múltiple

Primer Mito: Más armas, menos crímenes

John Lott, un economista de la Universidad de Yale, utilizó un modelo econométrico para argumentar que “el permitir a los ciudadanos llevar armas reduce los crímenes violentos, sin que aumenten las muertes por accidente”. Lott estimó que por cada un 1% de incremento en propietarios de

armas decrecía en un 3.3% en las tasas de homicidio de una población. Lott y su coautor, David Mustard, introdujeron la primera versión de su estudio en Internet en 1997, y unos miles de personas lo recogieron. Se convirtió en tema de debates políticos, de columnas periodísticas y con frecuencia de sofisticados debates en la Web. La discusión siguió las líneas ideológicas previsibles, con una destacada crítica que denunciaba el estudio por sus defectos metodológicos, incluso antes de disponer de una copia. En un libro con el título pegadizo *Más Armas, Menos Crimen*, Lott se mofó de sus críticos, acusándoles de anteponer la ideología a la ciencia.

El trabajo de Lott es un ejemplo de *ostentación estadística*. Dispone de más datos y de un análisis más complejo que cualquiera que estudie el tema. Obliga a quien quiera desafiar sus argumentos a sumergirse en un argumento estadístico muy complejo, basado en un conjunto de datos tan amplio, que ni siquiera puede manipularse con los ordenadores personales que la mayoría de los científicos sociales utilizan. Se alegra de compartir sus datos con cualquier investigador que lo desee, pero la mayoría de los científicos sociales se ha cansado de este juego. ¿Cuánto tiempo deben dedicar los investigadores a reproducir y criticar los estudios que utilizan métodos que han fracasado repetidamente? La mayoría de los investigadores sobre el control de armas ignoraron el reto de Lott y Mustard y siguieron con su trabajo. Dos reconocidos investigadores del ámbito de la Justicia criminal, Frank Zimring y Gordon Hawkins (1997: 57) escribieron un artículo diciendo que:

así como los Sres. Lott y Mustard pueden, con un modelo de determinantes del homicidio, producir datos estadísticos que sugieren que *la concesión obligada* (shall issue) reducirá el homicidio, esperamos que algún especialista en modelos estadísticos pueda realizar un tratamiento de los mismos períodos históricos con modelos diferentes y efectos opuestos. El modelo econométrico es un arma de doble filo en su capacidad para facilitar los resultados estadísticos que anima los corazones de los creyentes verdaderos de cualquier posición.

Zimring y Hawkins acertaron. Al año, dos especialistas, Dan Black y Daniel Nagin (1998) publicaron un estudio que demostraba que modificando el modelo estadístico un poco, o lo aplicándolo a diferentes subconjuntos de los datos, ya no se obtenían los resultados de Lott y Mustard. Black y Nagin encontraron que cuando Florida fue eliminada de la muestra no se encontró “ningún impacto relevante de las leyes del derecho a llevar armas en la tasa proporción de asesinatos y violaciones”. Llegaron a la conclusión de que la “inferencia basada en el modelo de Lott y Mustard era inadecuada y sus resultados no pueden ser utilizados para formular de forma responsable una *política pública*”.

Segundo Mito: Cuánta más gente sea encarcelada menos crímenes

El caso de Lott y Mustard fue excepcional, únicamente por la atención pública que despertó. Es muy común, incluso típico, que se publiquen estudios parecidos que utilizan métodos econométricos y que con el mismo conjunto de datos llegan a conclusiones opuestas. En una franca e inusual declaración de frustración con este estado de hechos, dos criminólogos, Thomas Marvell y Carlisle Moody (1997: 221), dieron a conocer un estudio que habían realizado sobre el efecto del encarcelamiento en las tasas de homicidios. Afirmaron que:

difundieron los [sus] resultados, junto con los datos utilizados, entre colegas especializados en el análisis cuantitativo. La contestación más frecuente fue desconfiar de los resultados, al margen de la rigurosidad del análisis estadístico. Lo que estaba detrás era la convicción, muy debatida informalmente pero pocas veces publicada formalmente, de que los científicos sociales, manipulando los procedimientos utilizados, pueden obtener cualquier resultado deseado. De hecho, la gran variedad de medidas relacionadas con el impacto de la población encarcelada se toma como la mejor prueba de la plasticidad de la investigación académica. La implicación, incluso entre aquellos que habitualmente publican estudios cuantitativos, es que no importa lo completo que sea el análisis, los resultados no serán creíbles si no son coherentes con las expectativas previas. Con un esquema de este tipo ninguna investigación puede tener éxito.

El gran mérito de Marvell y Moody fue reconocer abiertamente los problemas de la regresión múltiple y el haber hecho algunas sugerencias para mejorarla. Sin embargo, esta actitud es más la excepción que la regla entre los científicos, que a menudo se identifican tanto con sus modelos que se olvidan de la arbitrariedad de los mismos. Muchos de ellos llegan a creer que sus modelos son más reales y más válidos, que la propia realidad desordenada, recalcitrante e *incontrolada* que pretenden explicar.

Tercer mito: La pena capital reduce el crimen

En 1975 *The American Economic Review* publicó un artículo realizado por el economista, Isaac Ehrlich de la Universidad de Michigan, que estimaba que una ejecución evitaba ocho homicidios. Antes que Ehrlich, el especialista más conocido sobre en la efectividad de pena capital era Thors-ten Sellen, que había utilizado un método mucho más simple de análisis. Sellen diseñó gráficos para comparar las tendencias en los estados diferentes. No encontró diferencias, o muy pequeñas, entre los estados con o sin la pena capital, y llegó a la conclusión de que la pena capital no tenía ningún efecto. Ehrlich, en un acto de *ostentación estadística*, afirmaba que su análisis era más válido porque había controlado todos los factores que influyen en las tasas del homicidio.

Incluso antes de que fuera publicado, el trabajo de Ehrlich fue citado por el Fiscal General de los Estados Unidos en un informe jurídico (*amicus curiae*) ante el Tribunal Supremo de Estados Unidos en defensa de la pena capital. Afortunadamente, el Tribunal desestimó la evidencia de Ehrlich al no haber sido confirmada por otros investigadores. Una decisión acertada, puesto que al año otros investigadores publicaron un análisis estadístico, tan sofisticado como el de Ehrlich, que demostraba que la pena capital no tiene ningún efecto disuasorio

La polémica en torno al trabajo de Ehrlich fue tan importante que el *Consejo Nacional de Investigación* reunió a un grupo de expertos de gran prestigio para revisar el trabajo. Después de un estudio exhaustivo, el grupo afirmó que el problema no residía en el modelo de Ehrlich, sino en el uso de los métodos de regresión para resolver las cuestiones relacionadas con las políticas de justicia criminal. Llegaron a la conclusión (Manski, 1978: 422) de que:

dato que los datos disponibles en un análisis de ese tipo tienen limitaciones y puesto que la conducta delictiva puede ser muy compleja, no parece viable la existencia de un estudio conductual definitivo y concluyente, que permita zanjar el debate sobre los efectos conductuales de las políticas de disuasión.

A Ehrlich no le convencieron estos críticos y buscó las limitaciones de su trabajo. Sigue siendo el único creyente acérrimo de la validez de su modelo. En una reciente entrevista (Bonner y Fessendren, 2000) insistía en que "si se tienen en cuenta variables como el desempleo, la desigualdad de salarios, la predisposición y actitud favorable a utilizar la pena capital, ésta muestra un efecto disuasorio significativo".

Cuarto mito: La legalización del aborto redujo la criminalidad en los noventa

John Donohue y Steven Levitt publicaron en 1999 un estudio con una nueva explicación del marcado declive en los índices de criminalidad en los años noventa. Afirmaron que la legalización del aborto por el Tribunal Supremo americano en 1973 provocó una reducción del nacimiento de niños no deseados, un número desproporcionado que se habrían convertido en delincuentes. El problema con esto es que la legalización de aborto era un acontecimiento histórico concreto y existen pocos datos para un análisis de regresión válido. Es probable que los resultados varíen en función de cómo se seleccionen los datos para el análisis. En este caso, como indicó James Fox (2000: 303): "al utilizar solamente un indicador estadístico que resume el cambio en este periodo de doce años, [Donohue y Levitt] se pasa por alto la mayoría de los cambios en la delincuencia durante este período –la ten-

dencia ascendente durante el estallido de finales de los años 80 y la tendencia descendente en los años posteriores al estallido. Es como estudiar los efectos de fases de la luna en las aguas del mar pero recogiendo sólo datos de los períodos de bajamar”.

Cuando estaba escribiendo este artículo, incluí una frase que decía “pronto otro analista de la regresión volverá a analizar los mismos datos y probablemente llegará a conclusiones muy distintas” Unos días más tarde, apareció una noticia de prensa sobre un estudio de ese tipo. El autor era precisamente Lott de Yale, con John Whitley de la Universidad de Adelaida. Con los mismos datos llegaban a la conclusión de que “la legalización del aborto había provocado un aumento en las tasas de delincuencia, proximadamente entre un 0.5 y 7 %” (Lott y Whitley, 2001).

¿Por qué los resultados son tan dispares? Simplemente, cada grupo de autores eligen una manera diferente de analizar un conjunto inadecuado de datos. La econometría no puede convertir en una ley general válida el hecho histórico puntual de que el aborto se legalizara en los años setenta y que la delincuencia descendiera en los años noventa. Necesitaríamos al menos unas decenas de experiencias históricas parecidas para llegar a obtener una prueba estadística significativa.

Quinto mito: La reforma del Estado del Bienestar llevará a al pobreza a un millón de niños

El 1 de agosto de 1996, cuando el Senado de Estados Unidos planteó un cambio en las políticas de bienestar, el Instituto Urbano emitió un informe ampliamente difundido que venía a demostrar que: la reforma propuesta del Estado del Bienestar aumentaría la pobreza y reduciría los ingresos de las familias del grupo más bajo de ingresos... *Nosotros estimamos que 2.6 millones más de personas se situarán por debajo del umbral de la pobreza, incluyendo 1.1 millones de niños.* (El Instituto urbano, 1996, la pág., 1).

Los defensores del Bienestar apoyaron esta predicción, pero no convencieron a los diseñadores de la política. Los senadores que defendían la reforma simplemente no creyeron que los científicos sociales pudieran hacer predicciones válidas de este tipo. Y tenían razón. El Instituto Urbano no podría predecir la dirección del cambio, solamente su trascendencia. Después de la reforma de bienestar, la pobreza infantil se redujo, no aumentó. El modelo del Instituto Urbano era mucho más complejo que los otros modelos que hemos descrito en este artículo, pero la complejidad añadida sólo parece haber empeorado el problema.

Utilizando técnicas sofisticadas de *microsimulación*, recogieron correlaciones existentes previas, las introdujeron en ecuaciones complejas y

luego convirtieron estas ecuaciones en leyes generales. Toda su matemática se fundamentó en el supuesto de que si no se producía un cambio fundamental, la reforma de bienestar fracasaría. Todo el modelo se orientó a producir datos que avalaran sus argumentos y darles la apariencia de científicos. Pero la meta de la reforma era cambiar las cosas, y lo hizo.

Por qué fracasa la regresión

Aunque los modelos de regresión parecen complicados, en realidad son simplificaciones del mundo real. Para simplificar la matemática, la regresión utiliza las ecuaciones lineales. Esto significa suponer que si usted representa en un gráfico la relación entre dos variables la tendencia se parecerá a una línea recta. También presuponen que las variables se distribuyen según una curva normal clásica. Y suponen que los analistas diferencian entre las variables que son las causas y las que son los efectos. Evidentemente, los analistas de la regresión saben que el mundo real no encaja en estos presupuestos y, en compensación, hacen distintos ajustes en los datos. Pero los ajustes crean otros problemas. La única manera válida de probar un modelo después de todos estos ajustes es mostrar que funciona para predecir las tendencias futuras. Son ciencia basura aquellos modelos de regresión que no han demostrado funcionar con datos reales, distintos a los datos utilizados para ser originados.

Por qué fracasa la regresión: modelos lineales de un mundo no lineal

Cuando nos enfrentamos al mundo real (no lineal), el primer instinto del analista de la regresión es estandarizar y controlar los datos. Haciendo esto, se minimizan o eliminan los hechos históricos más interesantes e importantes. Terminan analizando un mundo estandarizado e idealizado que tiene poca relación con la realidad. Por ejemplo, consideremos las tendencias en el crecimiento de presos y los homicidios que Marvell y Moody (1997) querían explicar. Su artículo empieza con un gráfico que muestra las tendencias en presos por 100,000 habitantes y homicidios por 1,000,000 habitantes en los Estados Unidos. Es interesante y útil el gráfico que reproduce sus datos (ver gráfico 1).

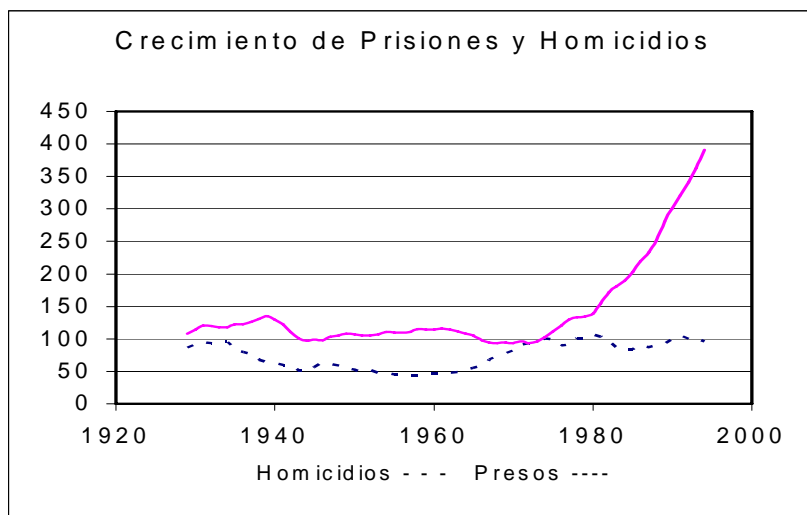
Lo más interesante de las tendencias que el gráfico representa son los puntos de inflexión, los lugares dónde las tendencias divergen de la linealidad. La tasa de homicidios aumentó enormemente entre mediados de los años sesenta y comienzo de los setenta, después se mantiene estable. El número de presos aumenta notablemente a comienzos de los años setenta, cuando en los Estados Unidos se construyeron más prisiones en respuesta

al aumento de la criminalidad. La tasa de homicidios estabilizada por los años ochenta, permaneció estable desde entonces.

En lugar de intentar explicar estos puntos importantes de cambio, Marvell y Moody utilizaron técnicas de regresión múltiple para *controlarlos*. Introdujeron controles para cada variable medible que se les ocurrió, incluyendo (Marvell y Moody, 1997: 209) “estructura de edad, factores económicos, subsidio de desempleo, raza, y variables indicadoras de la IIGM y la epidemia del estallido”.

Todos estos controles depuraron los cambios históricos más relevantes de sus datos. Esto les llevó a la conclusión de que un 10% de aumento en la población de presos suponía aproximadamente un 13% menos de homicidios. Sin embargo, si analizamos los datos de sus gráficos vemos que el declive esperado del 13% en la tasa de homicidios por cada 10% de incremento de presos desde 1975 sencillamente no se produce.

Gráfico 1



Aunque Marvell y Moody se encontraron con que la reducción esperada no se produjo, no les pareció razón suficiente para abandonar sus métodos estadísticos. Después de todo no estaban hablando del mundo real, sino de un mundo simplificado y purificado a través de una serie de ajustes matemáticos. Al enfrentarse a los hechos reales, afirmaron que si no hubiese aumentado el número de presos, la tasa de homicidios se hubiera agravado

mucho más. Y señalaron que, sin embargo, eso nunca habría pasado porque el gobierno habría tomado otras medidas para evitarlo.

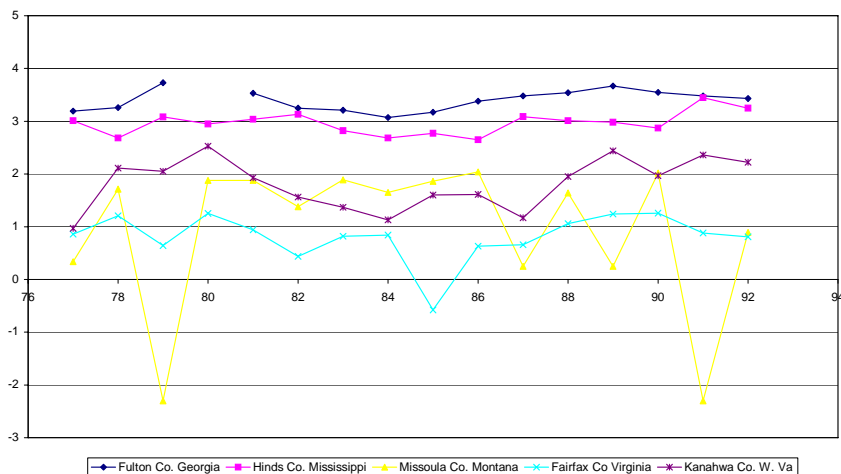
¿Pero qué valor tiene un análisis del que se extraen implicaciones que los propios autores señalan que nunca llegarían a darse? ¿Hasta qué punto es válida la teoría que subyace al análisis de la regresión múltiple si omite variables importantes, como las restricciones políticas, simplemente porque no pueden medirse cuantitativamente? ¿Cuánto dependen los resultados de las decisiones discrecionales sobre qué variables de control hay que introducir y cómo medirlas? Marvell y Moody se quedaron con estadísticas que les permitían decirnos lo que podría haber pasado si nada de que lo que realmente pasó hubiera pasado.

Por qué la fracasa la regresión: El mundo no es una curva normal

Además de la linealidad, la regresión múltiple presupone que cada variable está *normalmente* distribuida según el modelo de la curva en forma de campana. Es decir, que la mayoría de los casos deben agruparse en torno a la media dentro de cada categoría, con algunos casos en los extremos. Con frecuencia los datos contradicen este supuesto, de forma que llevan a resultados completamente erróneos. Un ejemplo representativo son los datos de John Lott sobre el control de armas.

Lott recogió un conjunto amplio de datos que facilitó generosamente a otros investigadores. Lamentablemente no empezó por la representación gráfica de sus datos, quizá porque tenía demasiados. Pero siempre es una idea buena empezar el análisis de datos con el diagrama para ver las tendencias antes de que se disimulen por todos los ajustes estadísticos. Si uno no puede utilizar todos los casos, merece la pena utilizar al menos los más representativos. Utilizando los datos de Lott, yo tracé la tendencia en las tasas de homicidio en varios estados donde la ley de *concesión obligada* había entrado en vigor durante el período cubierto por su estudio. Antes de aprobar las leyes, las autoridades locales tenían capacidad para conceder, o no, el permiso de armas. Después de que fueron aprobadas, tenían obligación de emitir el permiso de armas a cualquier adulto que lo solicitara. Si la hipótesis de Lott era correcta, debemos esperar que disminuya la tasa de asesinatos, una vez que se aprobaron las leyes.

El siguiente gráfico muestra las tendencias en la tasa de asesinatos de los condados principales en algunos estados, que adoptaron las leyes de *concesión obligada* entre 1977 y 1992. Las fechas en las que las leyes entraron en vigor varían de un estado a otro. Antes de seguir, el lector puede encontrar interesante examinar el gráfico e intentar inferir cuando la ley entró en vigor en cada condado (ver gráfico 2)

Gráfico 2**Tasas de homicidios en 5 Condados afectados por la legislación de armas**

Si se analiza el gráfico, se ve rápidamente que el patrón en el condado de Missoula, Montana, es completamente errático, con descensos muy pronunciados en las tasas de homicidios entre en 1979 y 1991. Sin embargo, no es un fenómeno real, sino el resultado de uno de los ajustes que Lott hizo para compensar para la no-normalidad de sus datos. En vez de utilizar los números reales, convirtió sus números a los logaritmos naturales. Algo muy habitual, puesto que los logaritmos naturales suelen ajustarse mejor que los datos reales a los supuestos de la regresión múltiple. El resultado para los datos de John Lott en el condado de Missoula de esos años fue igual a -2,30. Un dato curioso, porque que la tasa de homicidios de un condado no puede estar por debajo de cero, a menos que traigan de nuevo a la vida gente previamente asesinada. Sin embargo, no hay suerte. Para conseguir la tasa real de homicidios en cada condado, uno tiene que invertir los logaritmos de los datos de Lott con la fórmula $tasa\ real = e^{tasa\ logarítmica}$, donde $e=2.71828$. Algo que se puede hacer fácilmente con la función e^x de una calculadora científica. Apuntando -2,3 en la calculadora y pulsando e^x se obtiene .100, o un décimo de homicidios por 100.000 habitantes. El dato real para los homicidios en el condado de Missoula en 1979 y 1991 era cero. Lott utilizó “1” en vez de cero, porque el logaritmo natural de cero es

indefinido; dejarlo en cero significaba perder datos. Hay muchos -2.3 en su base de datos sobre homicidios, porque muchos de los condados son muy pequeños, mucho más que Missoula que contaba con 81.904 habitantes en 1992.

Al margen de las distorsiones en el condado de Missoula provocadas por la conversión de los datos a logaritmos naturales, las tendencias en estos condados son bastante planas. No hay un efecto claro de la introducción de las leyes de *concesión obligada* en el Condado de Missoula en 1991, en el Condado de Fulton (Atlanta, Georgia) en 1990, en el Condado de Hinds (Jackson, Mississippi) en 1990, en el Condado de Fairfax (Fairfax, Virginia) en 1988 y Condado de Kanawha (el Charlestown, Virginia occidental) en 1989.

Claro que podemos preguntarnos ¿por qué estamos analizando estos condados medianos en lugar de los centros de la población mayores? Esto me llevó a mi primera pista del error fundamental del argumento de Lott. Mi tendencia inicial fue representar las tendencias en las ciudades más grandes de América, porque es donde el problema del homicidio es muy grave. Pero rápidamente descubrí que en ninguna de estas ciudades existía la ley de *concesión obligada*. Estas leyes se aplicaron principalmente a los estados de poca población. Lo que significa que los datos de Lott no cumplen los supuestos fundamentales de un análisis de regresión. Para trabajar adecuadamente, la regresión múltiple exige que la variable “leyes de *concesión obligada*” se distribuya normalmente en el conjunto de datos. Los cálculos matemáticos que *controlan* las relaciones espurias no se pueden realizar si no existe un rango suficiente de variación en las variables importantes. Ésta era la *cortina de humo* del conjunto de tablas y las ecuaciones sofisticadas de Lott. En ningún lugar del libro reconoce este hecho. Cuando le pregunté por esto, se encogió de hombros. Para él no era un problema puesto que había controlado el tamaño de la población.

Estas irregularidades en los datos invalidan el análisis de Lott. Dos años antes Ayres y Donohue (1999) lo habían verificado en un análisis econométrico, pero Zimring y Hawkins pusieron a cero el problema de 1979. Habían estudiado la legislación del control de armas, sabían que las leyes de *concesión obligada* se implantaron en estados donde la Asociación Nacional del Rifle tenía mucha fuerza, sobre todo en el Sur, el Oeste y en las regiones rurales. En estos estados existían pocas restricciones sobre las armas. Afirieron que esta circunstancia legislativa invalida (Zimring y Hawkins 1997: 50) "nuestra capacidad para comparar las tendencias en estados *de concesión obligada* con las tendencias en otros estados. Dado que los estados que cambiaron la legislación son diferentes en la situación y constitución a los

estados que no lo hicieron, las comparaciones por las categorías legislativas corren el riesgo de confundir la influencia de los factores demográficos y regionales con el impacto conductual de sistemas legales diferentes". Zimring y Hawkins (1977: 51) a continuación afirmaban que:

Lott y Mustard eran conscientes de este problema. La solución que adoptaron, una técnica econométrica estándar, fue construir un modelo estadístico que controlara todas las diferencias entre las ciudades de Idaho y Nueva York que influyen en las tasas de homicidio y delincuencia, en lugar de hacerlo con las leyes de "concesión obligada". Si uno puede especificar las influencias principales en homicidios, violaciones, robos y hurtos en nuestro modelo, entonces puede eliminar la influencia de estos factores sobre las distintas tendencias. Lott y Mustard construyeron modelos que estimaban los efectos de los datos demográficos, datos económicos y sanciones penales de los distintos delitos. Estos modelos son el resultado del arte culinario estadístico en la medida que se crearon por los autores para este conjunto de datos y únicamente se probaron en aquellos datos que se utilizan en la evaluación de los efectos del derecho a llevar armas.

Lo que Lott y Mustard hicieron fue comparar las tendencias de Idaho y Virginia occidental y Mississippi con las tendencias en Washington y la ciudad de Nueva York. Lo que realmente pasó fue que por los años ochenta y comienzo de los noventa se produjo una explosión de homicidios en las principales ciudades orientales, en las que la mayoría de las personas iban bien armadas a pesar de la falta de permiso de armas. El argumento final de Lott se redujo a afirmar que en la mayoría de los estados rurales y occidentales de *concesión obligada* se evitó la epidemia de explosión de homicidios debido a sus leyes. Esto nunca se habría tomado en serio si no se hubiera disimulado con un complejo laberinto de ecuaciones.

Por qué fracasa la regresión: ausencia de comprobación predictiva

La prueba sólida en el análisis estadístico es la predicción. Un modelo útil debe hacer predicciones que sean mejores que la estimación aleatoria. Solamente de esta forma se puede distinguir entre causa y efecto y se pueden comprobar las predicciones causales. Los analistas de la regresión suelen hacer esto con datos históricos y utilizan datos de otro tiempo para predecir los más actuales. El problema con esto es que, cuando el resultado ya es conocido, es demasiado fácil ajustar el modelo para adecuarlo al resultado. Es como utilizar el tiempo de anteayer para predecir el tiempo de ayer, o utilizar las cotizaciones de anteayer para predecir las de ayer. Puede ser útil como forma de aprendizaje, pero la prueba realmente sólida es predecir el tiempo o la cotización de mañana. Este criterio, el éxito en la predicción,

se utiliza para evaluar los modelos de los mercados financieros, el tiempo, los resultados médicos, las tendencias de población y otros muchos otros fenómenos. Estos modelos funcionan de forma defectuosa y la regresión nos proporciona una buena medida precisamente de eso.

Desgraciadamente, los investigadores que utilizan técnicas econométricas para evaluar las políticas sociales generalmente no someten sus modelos a las pruebas de predicción. Podrían hacerlo, o bien hacer predicciones futuras y esperar a ver qué ocurre, o, si eso tardara mucho tiempo, entonces desarrollar su modelo con los datos de una población y luego utilizarlos para predecir los resultados de otra población. Sin embargo, los investigadores no hacen eso, o si lo hacen los modelos no funcionan de forma que los resultados nunca se publican (El Instituto Urbano hizo una predicción, pero no esperó a obtener los resultados para publicar sus conclusiones. Cuando los datos mostraron que su modelo no funcionaba, simplemente los retiraron de su página web.)

Las revistas que publican estos estudios no exigen pruebas de predicción, lo que indica que editores y críticos tienen unas expectativas bajas de sus campos. Al contrario, los investigadores recogen datos de un periodo determinado de tiempo y realizan ajustes y reajustes en sus modelos hasta que puede *explicar las tendencias que ya han ocurrido*. Existen muchas formas de hacerlo y con las computadoras modernas no es muy difícil seguir intentándolo hasta lograr que un ajuste funcione. En ese momento, el investigador se detiene, redacta los resultados y envía el artículo para que sea publicado. Más tarde, otro investigador puede ajustar el modelo para obtener un resultado diferente. Esto llena las páginas de las revistas de las ciencias sociales y permite a los profesores jóvenes conseguir la titulación académica. Nadie quiere reconocer que no se está haciendo ningún progreso.

La alternativa: volver a los gráficos y las tablas inteligibles

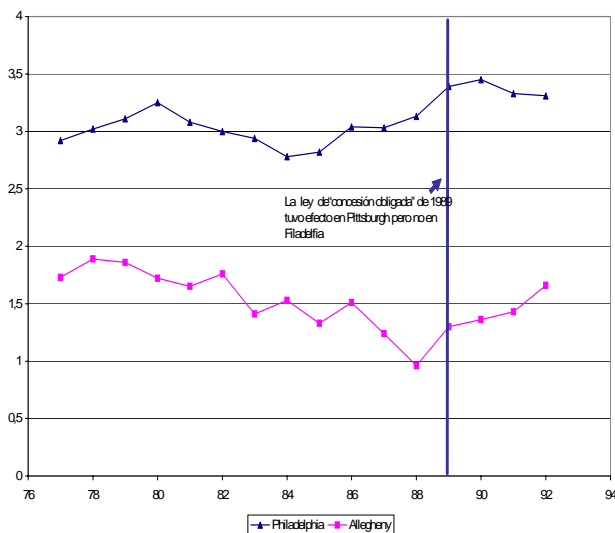
Cuando la ciencia basura se divulga en los medios de comunicación por los investigadores de universidades de gran prestigio y se publica en las revistas de gran impacto, lógicamente las personas se vuelven escépticos sobre el valor de investigación social. Hace unos años *The economist* (el 13 de mayo de 1995) publicó una editorial irónica que anunciaba que el *74.6% de sociología era basura*. Los más cínicos se preguntaron si la estimación no era algo baja. Pero es importante no tirar al bebé con el agua del baño. Existen trabajos serios en la sociología, la criminología y otras ciencias sociales, aunque no pueden publicarlos en las revistas que valoran la complejidad estadística más que la fiabilidad de los resultados. El trabajo más

fiable utiliza técnicas estadísticas más simples que no requieren tanto ajuste y estandarización de los datos. Tiene además la ventaja de que se puede leer y utilizar por las personas que no han consagrado años de sus vidas a aprender las técnicas de econométricas indiscifrables.

Los estudios que hacen un uso extensivo de los gráficos, como los de Sellin (1959) y Blumstein y Wallman (2000) han tenido mucho más éxito y han sido mucho más informativos que los estudios de regresión. Como ejemplo del poder de las técnicas de representación gráfica simple, podemos representar algunos de los datos sobre el control de armas que utilizó John Lott. Cuando un conjunto de datos es tan grande, quizá sea necesario seleccionar una parte pequeña para su representación gráfica, que será también muy informativa si la selección se hace bien. Repasando los datos de Lott, vi que esa condición se daba en el estado de Pennsylvania, la ley de *concesión obligada* se aprobó en 1989, salvo en la ciudad de Filadelfia. Era una oportunidad excelente para un *experimento natural*, puesto que se podían comparar las tendencias en dos áreas metropolitanas que se diferenciaban en una variable importante.

Gráfico 3

Tasas de homicidios en Philadelphia y Allegheny (incluye Pittsburgh)



El gráfico 3 compara las tendencias en Filadelfia, que es una ciudad y un condado, con las tendencias del Condado de Allegheny que incluye Pittsburgh. El gráfico muestra que las tasas de homicidios en general son

superiores en Filadelfia que en el Condado de Allegheny, pero el permiso de armas que la ley concede a los ciudadanos no muestra el efecto positivo señalado por John Lott. De hecho, la tasa de homicidios en el condado Allegheny había descendido antes de la aprobación de la ley y luego aumentó ligeramente. En Filadelfia, la tasa de homicidios había ido aumentando y después se niveló, a pesar de que la nueva ley no aplicó en esa ciudad. Las estadísticas de crímenes violentos muestran el mismo patrón en los dos condados analizados. Separando de esta forma los datos, podemos utilizar nuestro conocimiento cualitativo e histórico para interpretar las tendencias estadísticas. Para desacreditar este tipo de resultados, los defensores del permiso de armas tendrían que mostrar cómo otros factores compensaron el fracaso de la ley en tener un efecto claro.

Conclusiones

Estos casos pueden ser suficientes para convencer a la mayoría de los lectores que la regresión múltiple no es de mucha utilidad para demostrar relaciones causales, al menos en lo relacionado con el impacto real de las políticas sociales. De hecho, el problema es más amplio y muchos especialistas dudan que la regresión múltiple sea una manera válida de establecer cualquier argumento teórico. En 1985, Stanley Lieberman (1985: ix), un profesor distinguido de la Universidad de California, escribió “yo estoy completamente de acuerdo con las metas de la investigación empírica de gran parte de la sociología contemporánea, con su énfasis en el rigor y la cuantificación. Sin embargo, me cuesta aceptar que gran parte de los procedimientos y supuestos de esta empresa no tienen mucho más mérito que la búsqueda de una máquina de movimiento perceptivo”. En 1991, David Freedman, un conocido sociólogo de la universidad de California en Berkeley y autor de manuales sobre métodos cuantitativos de investigación, conmocionó los fundamentos de los modelos de regresión en las ciencias sociales cuando declaró “francamente yo no creo que la regresión pueda ser muy relevante en una relación causal. Ni las ecuaciones de regresión, por sí mismas, nos sirven mucho para controlar variables extrañas” (Freedman, 1991: 292).

El artículo de Freedman provocó una fuerte reacción. Richard Berk (1991: 315) señaló que la afirmación de Freedman “sería muy difícil de aceptar por la mayoría de los sociólogos cuantitativos”. Llega al corazón de su empresa empírica y, al hacerlo, pone en riesgo sus carreras profesionales.

Las ciencias sociales no disponen de procedimientos adecuados para reconocer el fracaso de un método de investigación ampliamente utilizado.

Métodos que, incorporados en las licenciaturas de las principales universidades y publicados en las revistas de prestigio, tienden a perpetuarse. Muchas personas creen que si un estudio se publica en una buena revista existe garantía de su validez. Los casos que hemos analizado demuestran que no es así. Los críticos aseguran que se siguen las prácticas establecidas, pero esto sirve de poco cuando las mismas prácticas son defectuosas.

Encontrar los fallos en los estudios de la regresión es difícil. Normalmente, la única manera de estar seguro es conseguir los datos originales y analizarlos de nuevo. Esto es lo que cabe esperar de un crítico profesional. Lleva tiempo reproducir un estudio de regresión múltiples, normalmente un o dos años, y muchos estudios nunca se reproducen porque no todos tienen un interés general. Incluso cuando se realiza una repetición y no se confirma un estudio, los editores de la revista pueden pensar que es un simple caso de debate científico normal. Sin embargo, el problema es que no se produce ningún avance. No estamos más cerca de disponer de un modelo matemático útil para predecir la tasa de homicidios que cuando en 1975 Ehrlich publicó su artículo.

No existen resultados importantes en sociología o en criminología que sean tan complejos para que no puedan difundirse mediante gráficos y tablas que son fáciles de comprender por cualquier persona inteligente y por los políticos. “Ha llegado el momento de admitir que el emperador no lleva ropa”. La regresión múltiple y otras técnicas matemáticas sencillamente no han cumplido su promesa de ser un método para evaluar el impacto de las políticas públicas. Los estudios que afirman lo contrario son ciencia basura.

References

- Ayres, I.-Donohue, J. (1999): Nondiscretionary Concealed Weapons Laws: A Case Study of Statistics, Standards of Proof, and Public Policy. *Am. Law and Ec. Rev* 1: 436-470.
- Berk, R.A. (1991): Toward a Methodology for Mere Mortals, *Sociological Methodology* 21: 315-324.
- Black, D.-Nagin, D. (1998): Do Right-to-Carry Laws Deter Violent Crime? *J. Legal Studies* 27: 209-219.
- Blumstein, A.-Wallman, J. (eds.) (2000): *The Crime Drop in America*, Cambridge University Press, New York, pp. 13-44.
- Boldt, D. (1999): Study Evidence on Guns, *Philadelphia Inquirer*, Dec. 14. Downloaded on May 17, 2000 from <http://www.phillynews.com/inquirer/99/Dec/14/opinion/BOLDT14.htm>.
- Bonner, R.-Fessenden, F. (2000): States with No Death Penalty Share Lower Homicide Rates, *New York Times*, September 22. Downloaded from: <http://www.nytimes.com/2000/09/22/national/22DEAT.html>.

- Donohue, J.-Levitt, S. (1999): Legalized Abortion and Crime. Stanford University Law School. Downloaded in August, 2000 from: [http://papers.ssrn.com/paper.taf? ABSTRACT_ID=174508](http://papers.ssrn.com/paper.taf?ABSTRACT_ID=174508).
- Ehrlich, I. (1975): The Deterrent Effect of Capital Punishment: A Question of Life and Death, *Am. Ec. R.* 65: 397-417.
- Fox, J. (2000): Demographics and U.S. Homicide. En A. Blumstein, J. Wallman (eds.), *The Crime Drop in America*. Cambridge University Press, New York, pp. 288-317.
- Freedman, D. A. (1991): Statistical Models and Shoe Leather. *Sociological Methodology* 21: 291-313.
- Lott, J.-Mustard, D. (1997): Crime, Deterrence and the Right to Carry Concealed Handguns, *J. Legal St.* 26: 1-68. Downloaded on August 10, 2000 from: <http://www.journals.uchicago.edu/JLS/lott.pdf>
- Lott, J. (2000): *More Guns, Less Crime: Understanding Crime and Gun Control Laws*. University of Chicago Press, second edition with additional analyses.
- Lott, J.-Whitley, J. (2001): Abortion and Crime: Unwanted Children and Out-of-Wedlock Births, Yale Law & Economics Research Paper No. 254. Downloaded on July 9, 2001 from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=270126.
- Manski, C. (1978): Prospects for Inference on Deterrence through Empirical Analysis of Individual Criminal Behavior. En Blumstein, C. et. al, eds, *Deterrence and Incapacitation*, Washington: National Academy of Sciences, pp. 400-424.
- Marvell, T.-Moody, C. (1997): The Impact of Prison Growth on Homicide. *Homicide St.* 1: 205-233.
- Sellin, T. (1959): *The Death Penalty*. American Law Institute, Philadelphia.
- Zimring, F.-Hawkins, G. (1997): Concealed Handguns: The Counterfeit Deterrent, *The Responsive Community* 7: 46-60.

Ted Goertzel es Profesor de Sociología en la Universidad Rutgers en Camden, New Jersey, USA. Es autor de una biografía del sociólogo y Presidente de Brasil, Fernando Henrique Cardoso, que fue publicado por Lynne Rienner en los EEUU y que se publicará en portugués por Saraiva Ed. de Sao Paulo. Recientemente ha publicado "The World Trade Center Bombing as a Fourth Generational Turning Point", y participa en distintos cursos sobre Ciberespacio y Sociedad, Métodos y técnicas de investigación social y Ética y política en la Justicia penal.
Department of Sociology. Rutgers University. Camden, NJ 08102. USA