Westlaw Today
powered by Reuters

Why curve-fitting cannot be used to show causation or estimate impact

By Iain M. Cockburn, Ph.D., Boston University

AUGUST 17, 2021

For civil litigation to "work" in the sense of giving injured parties adequate compensation, and neither over- nor under-penalizing bad actors, expert opinions must be based on reliable methods. Courts are often put in the position of assessing statistical analyses purporting to establish the causal impact of some challenged conduct, or to quantify its impact.

But while much progress has been made in economics and other academic disciplines to set a high standard for academic researchers to make causal claims, in my experience as a testifying expert, inappropriate statistical methodology keeps creeping back into our courtrooms, notwithstanding doctrines like *Daubert*, and the best efforts of judges to exclude "junk science."

One such methodology is curve-fitting, which can be particularly dangerous in the hands of an unscrupulous or naïve user.[1] To see why, imagine the role of an academic health economist retained to provide an expert opinion on damages in a case involving allegedly improper marketing of a pharmaceutical product.

Assuming that liability is found, a key input in calculating damages would be the difference between actual sales of the product and an estimate of how much would have sold but for the alleged improper marketing.

The court is likely to require convincing evidence of the causal impact of promotion on prescribing to estimate the but-for sales — that is to say, the amount by which sales would increase or decrease if, all else equal, the challenged promotion had not taken place.

The problem is that economists have cautioned for at least the last 50 years that marketing and sales are "endogenous variables," that is, jointly determined by market opportunities and the seller's response to them, or subject to reverse causality (e.g., this year's marketing budget depends on last year's sales). This makes the causal effect of marketing quite difficult to identify.

For instance, suppose a drug receives approval for use in treating an important new indication. The manufacturer is then likely to increase its promotion activity to try to heighten awareness and encourage physicians to prescribe it, at the same time that physicians may be changing their prescribing patterns based on other sources of information.

Did the increased marketing cause physicians to prescribe the drug more frequently for this indication? Or were they responding quite independently to the publication in a professional journal of the results of a clinical study completed to obtain the FDA's approval of the new indication, or word-of-mouth among their colleagues, or the relaxation of formulary constraints on using the drug in this way? Or all of the above? Or was it the increase in prescribing of this drug that led marketing managers to "follow the money" and reallocate sales force effort? These questions are difficult to answer, particularly in aggregate market data where sales volumes and marketing expenditures tend to move together over time.

Inappropriate statistical methodology keeps creeping back into our courtrooms, notwithstanding doctrines like Daubert, and the best efforts of judges to exclude "junk science."

Sometimes there is sufficient independent variation in the prescribing patterns of individual physicians, or across geographic locations, to try to identify the impact of marketing while controlling for other factors. This often leads economists to look to disaggregate data to try to estimate causal effects. But unfortunately, endogeneity problems are generally still present.

For example, physicians are often targeted for marketing based on their likelihood of prescribing. Cardiologists are more likely to prescribe heart drugs than non-cardiologists and are thus more likely to be "detailed" by sales representatives of the manufacturer of a heart drug. Observing more prescriptions for this drug being written by detailed physicians versus non-detailed physicians thus tells you nothing about whether the detailing caused the prescribing.

A carefully designed study that uses data on prescribing over time by individual physicians might be able to address this type of problem, since it will be able to control for the medical specialty of the prescriber.

But even looking at the same prescriber over time can be problematic — for example, a prescriber may switch employment and start being detailed because now the prescriber has been

identified as working in a relevant specialty, and at the same time starts prescribing more of the drugs used for treating the patients in the new practice.

Did the detailing cause the physician's change in prescribing, or was it changing employers?

Or suppose the analysis is based on comparing marketing and prescribing across different counties.

Opening a new health facility in a county would likely lead to both more prescriptions written for residents of that county, as patients gained access to more convenient care, and to more marketing effort in that county as the local sales force saw expanded opportunities to detail physicians.

In both of these situations, these events would contribute to an observed association between sales and marketing but do not reflect a causal relationship.

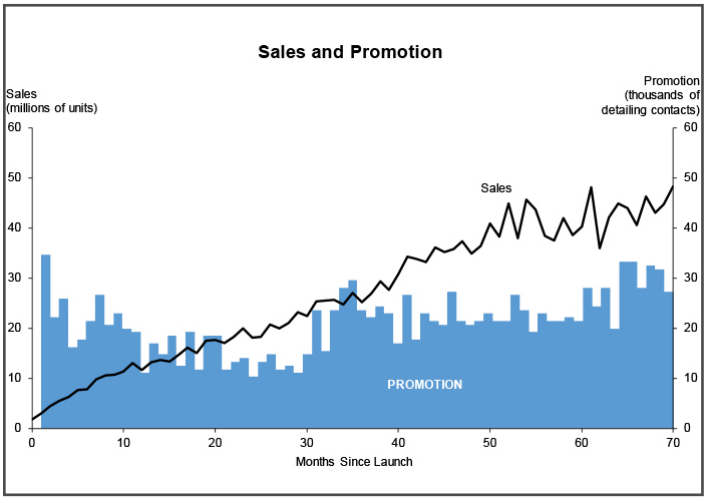
It is difficult to get this right, but it is also critically important. Sloppy analysis, or misinterpretation of its results, can lead to completely incorrect conclusions.

Identification of causal effects has become a central concern of the empirical literature in economics and other social sciences, to such an extent that professional gatekeepers such as journal editors or promotion committees have been dubbed "the identification police."

Getting a publishable estimate of the causal effect of marketing that meets professional standards and passes peer review normally requires one to use advanced econometric techniques, or a study design based on randomization of "treatment" or otherwise resembling a natural experiment.

Unfortunately, in my experience the endogeneity problem is often disregarded in the courtroom setting. Indeed, some experts not only ignore that "correlation is not causation," but also go to great lengths to "find" a correlation when in fact very little true association exists.

To do so, they typically start by collecting some historical data on sales and promotion. For example, the historical data on sales and promotion for the prescription medication Tagamet since launch is charted below.[2]



The first problem an expert trying to establish a correlation between promotion and sales for Tagamet would have to solve is that, while sales keep increasing, promotion goes up and down — if anything the overall trend over time in promotion is pretty flat. If one is going to be able to report a strong correlation between sales and promotion, one will need to find a way for them to track more closely.

Here, the peer-reviewed literature in economics and marketing can actually help: it is generally recognized that where marketing does influence purchase decisions, the effect is somewhat persistent — this month's purchases can reflect the cumulative impact of past marketing efforts that have built up "brand equity" or communicated messages that have not been forgotten.

Identification of causal effects has become a central concern of the empirical literature in economics and other social sciences.

Indeed, academic studies of drug promotion have often constructed a "stock" of promotion built up by cumulating each month's promotion over time while allowing for some "depreciation" or forgetting.

It will help the argument that promotion is responsible for sales if this stock of promotion "fits" the sales data well. The stock of promotion is constructed from monthly data on the number of sales calls, cumulated over time and net of depreciation.

Here is where the first econometric rabbit goes into the hat: depending on the depreciation rate that is used, the stock of promotion will grow faster or slower than sales, or — best of all — grow at approximately the same rate.

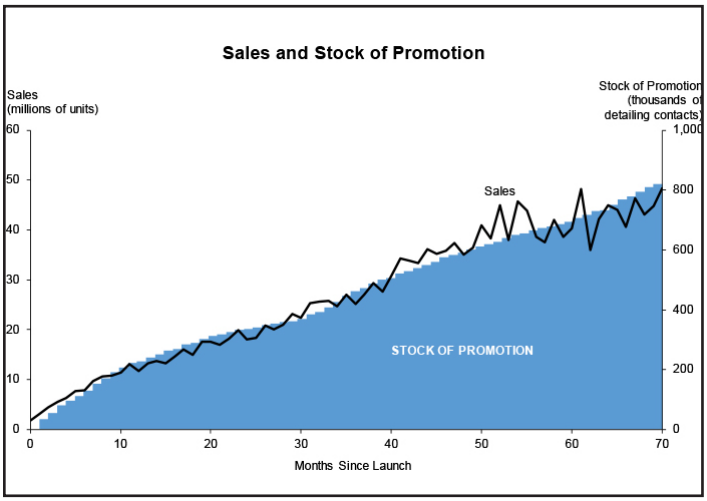
What is the appropriate depreciation rate to use? It would be possible to find some independent estimates from the peer-reviewed literature. But these are mostly large enough that when applied to the data on monthly promotion the result is a stock of promotion that does not grow fast enough to track the growth in sales, or even declines.

An alternative approach is to "estimate the depreciation parameter from the data." This will result in the best possible tracking of sales. One can select the depreciation rate that transforms the monthly promotion figures into a steadily rising stock that will track sales.

This looks promising. Now one can do a regression analysis, which has the advantage that one can include other "explanatory variables" in the model which "account for" other things that an opposing expert may try to argue were driving sales.

The regression model will have a large, positive and statistically significant estimated "coefficient" on promotion. Provided one is willing to ignore professional standards in economics, one

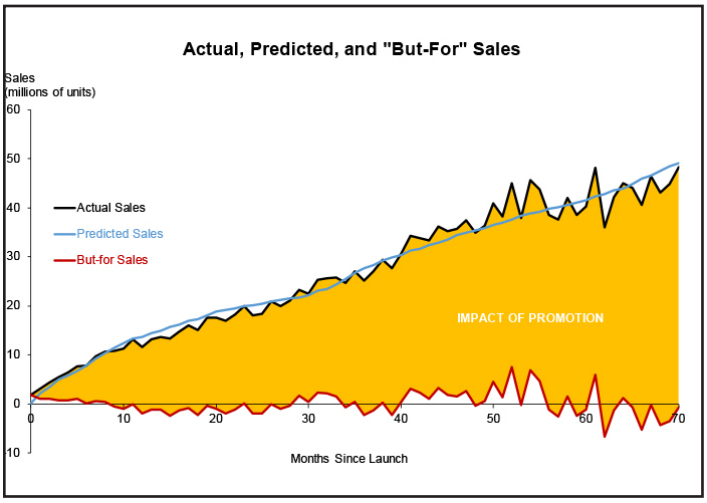
can now say in an expert report that "econometric analysis" has "demonstrated a causal relationship between promotion and prescribing."



For these data, if you assume that the promotion coefficient establishes the magnitude of a causal relationship between promotion and sales, then the model can be used to make a counterfactual but-for prediction as to what sales would have been with different amounts of promotion.

Here the model implies that without promotion, sales would have been close to zero. This is shown graphically in the chart below where the but-for sales (i.e., the sales predicted by the model when promotion is set to zero) hover around zero.

The difference between actual and but-for sales (the yellow area in the chart) is very large. So if the promotion is found to violate some statute, or is otherwise improper, the model will support a very large claim for damages.

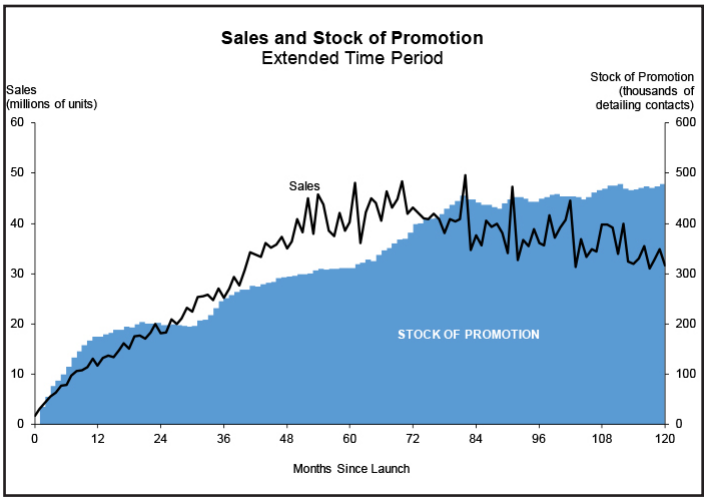


An opposing expert might point out that the regression model shows no such thing, for any number of technical reasons; that at best there appears to be an association not a causal effect; and that the interpretation of the promotion coefficient as the size of the causal effect would never be accepted by a peer-reviewed journal.

Indeed, the size of the estimated promotion effect is nothing like estimates found in the peer-reviewed literature.[3] But one can counter that argument by telling the court that the model should be trusted as it does an excellent job predicting actual world sales, in the sense that the sales estimates that result from applying the model's regression coefficients to actual levels of promotion and other explanatory variables are close to actual sales.

This can be seen in the chart above by comparing the predicted sales line with the actual sales line. Indeed, the model's "R^2" — the ratio of predicted variation to actual variation in sales — is satisfyingly high, with the model "explaining" 96% of the variation in sales over time.[4]

So far, so good. But what is an expert to do if at some point the trend in sales reverses and prescriptions start to decline, as happened for Tagamet about six years after launch? As shown below, in this case, even with the best-fit depreciation rate, stock of promotion and sales will no longer track each other so closely.

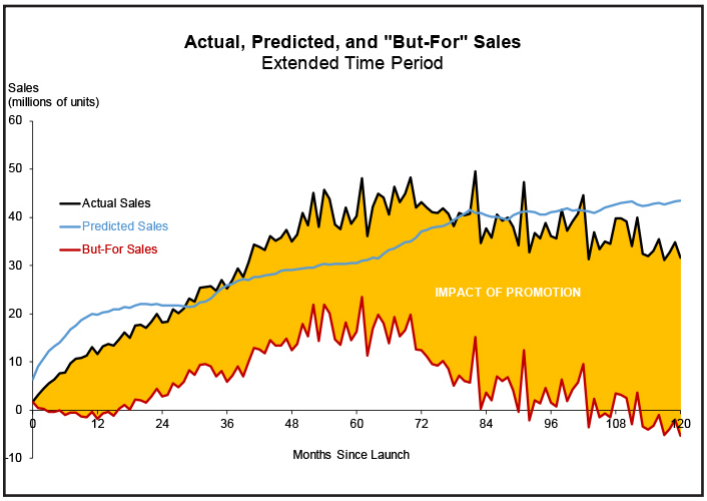


The problem is that after month 70 or so, the promotion and sales variables are trending in opposite directions. As a result, the model no longer implies that, without promotion, sales would have been close to zero. Moreover, as shown in the chart below, predicted sales based on the model look very different from actual sales.

Now the model badly overpredicts sales for the first two years, underpredicts for the next five, and then predicts steadily growing sales over the remainder of the period while actual sales declined. This will cause even a statistically unsophisticated observer to question the validity of the model.

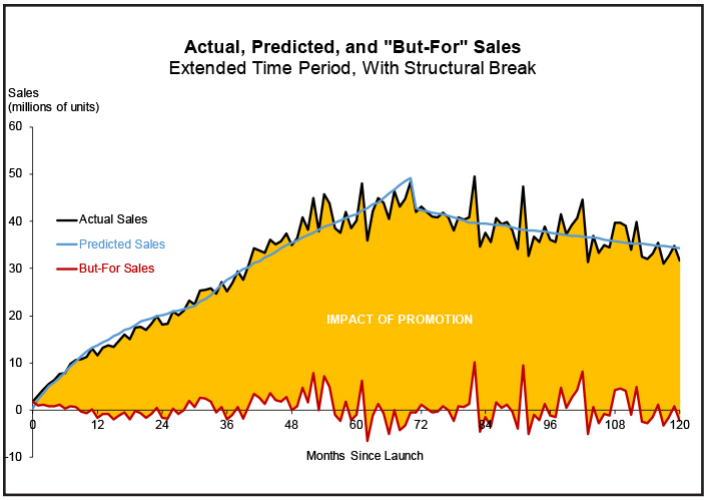
Here is where a second econometric rabbit goes into the hat! One can easily make this problem go away by adding a "structural

break" to the model to account for unobserved or hard-to-observe factors that change the relationship between marketing and sales after the 70th month.



One can justify this by appealing to "different X environments pre- and post-month 70," where X could be "economic" or "regulatory" or "professional opinion" or "guidelines" or "third-party payor requirements," or anything else that sounds plausible.

With separate effects of marketing pre- and post-month 70, one can get a much better-fitting model, where sales predicted by the model once again track actual sales closely.[5]



Now the model is back to predicting that, but for promotion, sales would have been close to zero.

Should a court accept an analysis like this? In my opinion — which is consistent with that of the vast majority of economists and statisticians — it should not.

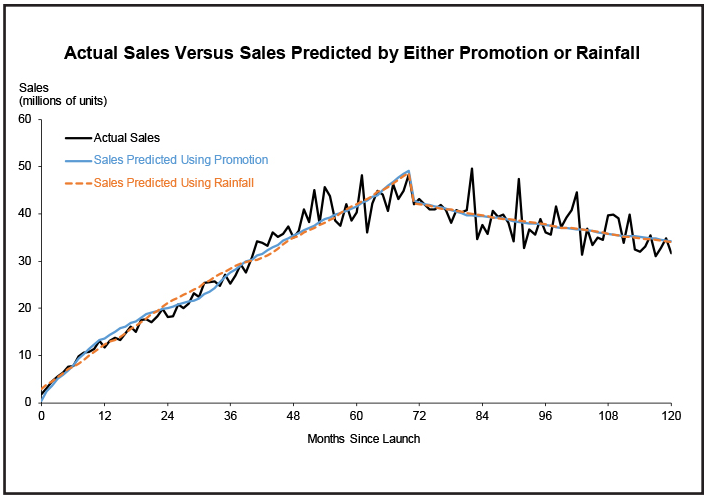
This type of curve-fitting does not, indeed cannot, demonstrate the presence of a causal effect or reliably estimate its magnitude, and is therefore a very poor predictor of how things would have played out in a counterfactual world in which the promotion had not happened.

One way to see how the constructed model is invalid — and utterly inappropriate for making but-for projections of sales, or establishing a causal relationship between actual outcomes and the allegedly improper promotion — is to recognize that the entire analysis presented here will generate essentially identical results using pretty much any series of data that has non-negative values to "explain" sales.

For example, as shown below, a model that predicts sales using monthly inches of rainfall in the U.S. in place of promotion generates almost identical results.[6] That is, such a model "fits" the data extremely well, with an R^2 of over 99%, and generates a prediction of almost zero sales if we assume no rain over the period in question. But obviously rainfall is not driving Tagamet sales.

Indeed, some experts not only ignore that "correlation is not causation," but also go to great lengths to "find" a correlation when in fact very little true association exists.

Not surprisingly, there is in fact no correlation between monthly rainfall and sales of this drug, and it is hard to come up with a plausible theory of why there should be such a correlation. Clever use of depreciation rates and structural breaks, however, can generate such an apparent relationship out of thin air.



This type of faulty analysis and concomitant wildly overstated damages estimates can be prevented by requiring that expert opinions conform to the set of practices and standards developed by economists and statisticians that constitute methodological safeguards to prevent this kind of erroneous inference.[7] Courts should pay close attention to them in deciding whether to accept this kind of testimony.

Notes

[1] Readers may recall the rosy forecast of COVID-19 deaths prepared in April 2020 by a White House economist who fit a curve with no basis in epidemiological principles to daily data on COVID-19 deaths, and "predicted" daily deaths in the U.S. peaking around April 21, 2020, and then falling to zero within a couple of weeks.

[2] The data are from figures in Ernst Berndt et al., Information, Marketing, and Pricing in the U.S. Antiulcer Drug Market, 85 Am. Econ. Rev. 100 (1995), who make no causal claims, and are careful to impose economic structure on their modeling of market shares.

[3] The bulk of well-conducted studies of the relationship between marketing and drug prescribing by researchers in economics and marketing have found that the impact of marketing is modest at best, with multiple doctor visits required to generate a single new prescription. Thus, a finding that a substantial volume of prescriptions was the result of marketing should also ring alarm bells.

[4] Econometricians generally caution against using R2 to evaluate a model's reliability, but it can have a powerful intuitive appeal to non-statisticians. Can the model be wrong if it does such a good job predicting real-world outcomes?

[5] Separate effects of promotion on sales are achieved by an introduction of a dummy variable for post-month 70 and its interaction with the stock of promotion into the model specification. An alternative way of achieving the same effect is to introduce a time trend for the post-month 70 period and its interaction with the stock of promotion.

[6] I could equally have chosen points scored by my favorite sports team, tons of gold held by the Federal Reserve, or the number of babies born in Belgium.

[7] The curve-fitting exercise documented here is particularly vulnerable to misuse or misinterpretation because it is not constrained by any kind of economic model of decision-making by the actors who generate market outcomes, and has no plausible source of identifying exogenous variation.

About the author

As the Richard C. Shipley Professor in Management at **Boston University**'s Questrom School of Business, **Iain Cockburn** specializes in competition and innovation in the life sciences, software and information technology industries. He has provided expert testimony in intellectual property, breach of contract, product misrepresentation and antitrust matters, addressing class certification, merits and damages issues. He can be reached at cockburn@bu.edu.

This article was first published on Westlaw Today on August 17, 2021.