Sub-Regressions in Antitrust Class Certification Can Be Unreliable

Kenneth Flamm[1]
Kenneth Flamm Economic Consulting
Austin, Texas
kflamm@gmail.com

Michael Naaman
Laurits R. Christensen Associates, Inc.
Madison, Wisconsin
mtnaaman@LRCA.com

Statistical analysis plays an important role in the litigation process, and in antitrust class actions in particular. Courts -- judges, juries, attorneys and other participants -- rely on experts to boil down sometimes complex statistical analysis into a form that can be understood by the non-statistician. Most importantly, courts are gatekeepers to the reliability of the statistical analysis presented. Therefore, to protect the integrity of legal proceedings and the usefulness of statistical analysis as a tool in those proceedings, it is imperative that the community of statistical experts call out the misapplication and misinterpretation of statistical analysis.

Unfortunately, a recent article in a legal forum does just the opposite. In **Sub-Regressions: A Rigorous Test for Antitrust Class Certification**, Laila Haider and Muneeza Alam uncritically promote what have been labelled "sub-regressions," arguing "it is critical to not junk a scientific and valid

[1] Dr. Flamm was an expert witness on behalf of plaintiffs in the Optical Disk Drive price-fixing class action. Dr. Naaman and colleagues at Christensen Associates provided analytical support to plaintiffs.

statistical approach that may be crucial for the economic analysis of class certification."[2] The Haider and Alam article does a good job documenting the mixed acceptance of sub-regressions across courts, which underscores the need for objective resolution of the issue. However, the title and slant of the article, placed in a well-read legal forum, may result in additional courts innocently accepting sub-regressions as statistically valid in situations where they are not. Particularly misleading in the article's title is the use of the term "rigorous," which Webster's Dictionary defines as "scrupulously accurate." This description simply cannot go unchallenged. To the contrary, as we show in this note, sub-regressions will often be statistically unreliable when used in small subsamples in antitrust class certification.

In general, the linear regression models used by economic experts in antitrust class certification should produce estimates of parameters in statistical models that have, at a minimum, the properties of *consistency* and *asymptotic normality*.[3] This means, respectively, that (a) as sample size grows large, parameter estimates approach closer to the true underlying value for the parameter; and (b) uncertainty about the true value of an estimated parameter can be measured, and approximated in large samples using the normal distribution (the bell-shaped curve that is the basis for most statistical inference). Both properties are *large sample properties*, however, and require reasonably large amounts of data in order to be relevant. Statistical estimators and inference relying on these large sample properties will often give inaccurate and incorrect answers when used in too small samples.

---

[2] Laila Haider and Muneeza Alam, "Sub-Regressions: A Rigorous Test for Antitrust Class Certification," *Law360*, December 5, 2014 (http://www.law360.com/articles/601614/sub-regressions-a-rigorous-test-for-antitrust-class-cert-?article_related_content=1).

[3] If the statistical disturbance terms in a regression equation are independent and drawn from a single common normal distribution (assumptions that can and should be tested, if used), then estimated parameters in a linear regression model can be reliably estimated, and uncertainty about their values precisely bounded even in a small sample. But economic theory rarely supports these assumptions, and often contradicts them. For example, price data, the most ubiquitous data used in class certification, cannot be produced by an economic model with truly normal statistical disturbance terms, since if this were the case, one would observe transactions at negative prices (i.e., sellers paying customers to take their products) some of the time, in violation of both economic theory and empirical observation.

It is easy to illustrate how slicing and dicing datasets into tiny subgroups (*i.e.,* sub-regressions) violates the theoretical assumptions of large samples underlying the statistical techniques typically utilized in class action econometric analyses, and can produce specious evidence of lack of injury within small subsets of class members, or subgroups of producers and consumers, when in fact there is actually an impact from the alleged illegal conduct on all class members. In these circumstances, rather than really being evidence of individual issues (lack of class cohesion), this outcome is actually evidence that the individual sample subsets used to estimate key parameters in sub-regressions are too small to produce anything but meaningless noise.

As one prominent econometrics textbook states, "…virtually all economists agree that **consistency** is a minimal requirement for an estimator. The Nobel prize-winning econometrician Clive W. Granger once remarked, 'If you can't get it right as *n* [sample size] goes to infinity, you shouldn't be in this business.' The implication is that, if your estimator of a particular population parameter is not consistent, then you are wasting your time."[4] Unfortunately, the sub-regression techniques described by Haider and Alam have been used in several of the cited cases with subsamples that are clearly too small, with the result being economically nonsensical findings trumpeted as evidence that the conduct lacked class-wide impact.

Nowhere in the Haider and Alam article is this *large-sample only* limitation on reliable use of these statistical techniques noted. Without careful observation of this limitation, spurious, fallacious "evidence" of lack of injury to all or nearly all class members can be generated in virtually any class action, by simply reducing the size of the subsamples in which sub-regressions are estimated. At best, such a slicing and dicing approach reflects ignorance of statistical properties; at worst it is statistical trickery passed off as "rigor." Regardless, it must be called out as invalid and unscientific.

[4] Jeffrey M. Wooldridge, *Introductory Econometrics: A Modern Approach*, (Cengage South-Western, 2013), p. 169.

This fundamental flaw of the sub-regression approach is easily illustrated using a simple example. Imagine there is a commercial mint producing coins in a large variety of denominations, colors, and sizes for a variety of nations scattered around the globe. There are 10,000 different coins being produced by the mint. The issue is whether coins using the mint's production techniques are perfectly balanced, producing an equal probability of heads and tails when flipped. Let us also suppose that we are allowed to randomly select one of each of these 10,000 coins and then we flip it twice in order to test its balance. We assign a counter variable a value of 1 if we get a head, and a zero if we get a tail. For a perfectly balanced coin, if we normalize this counter by dividing by the number of flips, we would expect this counter to approach a value 0.5 as the number of flips increases.

An incorrect way to test the hypothesis that all coins are equally and perfectly balanced is, after flipping each coin twice, to add the flips in a single counter variable *for each type of coin*, and see what values the counter variable takes on. The problem with this method is that even if all coins are perfectly balanced (homogeneity across coins), so the counters for all coins, individually and together, have an expected value of 0.5, with only 2 flips per coin type, approximately 25% of the individual coin counters in this sample of 10,000 coins will have value zero (corresponding to two tails), 25% will have value one (corresponding to two heads), while only 50% of the coins will produce a counter with a value of 0.5, the true value. A less than careful interpretation of the coin-by-coin test would incorrectly conclude from this outcome that there is heterogeneity in balance within this set of coins (with ¼ producing only heads, ¼ only tails, and only half fairly balanced) when, in fact, all coins are perfectly balanced. The method just described is in essence what is going on with the sub-regression method when it slices data into ever tinier chunks.

A vivid illustration of how sub-regressions can be employed as statistical trickery is available through a simple simulation we have created with an artificially generated dataset that imposes the exact same effect on all units within the class as the undisputed truth. Using the sub-regressions approach of slicing and dicing data into small groups of observations, carried out to the point where the required large sample assumptions underlying the statistical modeling technique no longer hold, yields nonsense results. Our simulation results greatly resemble results using actual datasets in at least some of the recent cases that the Haider and Alam article cites.[5] By using sub-regression techniques that rely on large samples for their validity and reliability, and instead applying them to very small samples, false "evidence" of heterogeneity is produced. That is, the touted sub-regression techniques reach the conclusion of different results for different groups even when it is known with perfect certainty that the economic parameter being estimated is the same across all groups.

In this simulation, designed to mimic a real dataset in a real class action certification case, we created an artificial dataset containing costs for 1000 different products, over anywhere from 12 to 48 time periods.[6] In this artificial dataset, we know with certainty that the true, underlying pass-through rate (the extent to which a change in costs is reflected in price) is exactly 100% and identical for all products. We can estimate a pass-through rate using a product-by-product version of the sub-regression methodology, and contrast it with using methodology that groups products together in order to estimate a single pass-through rate. We can then assess whether each of these different approaches is recovering what we know with certainty to be the correct answer—100% pass-through.

With 24 time periods, and three changes in costs distributed across them, for example, the results show the product-by-product pass-through sub-regression method finds positive and statistically significant pass-through rates in only 21.5% of the 1000 products! The other 78.5% of individual product pass-through rates estimated using the sub-regression method are statistically insignificant or even negative, even though we know with absolute certainty that the true pass-through rate is 100%. This

[5] See redacted "Declaration of Dr. Kenneth Flamm in Further Support of Indirect Purchaser Plaintiffs' Motion for Class Certification," *In re* Optical Disk Drive Prods. Antitrust Litig., No. 10-md-2143 (N.D. Cal. Mar. 7, 2014), ECF no. 1161, pp. 63-67, and Exhibit 3A.

[6] The costs are randomly generated with initial values that vary randomly by product, but the costs used in this simulation behave like actual values for real products in a real dataset. See redacted "Declaration of Dr. Kenneth Flamm in Further Support of Indirect Purchaser Plaintiffs' Motion for Class Certification," *In re* Optical Disk Drive Prods. Antitrust Litig., No. 10-md-2143 (N.D. Cal. Mar. 7, 2014), ECF no. 1161, pp. 63-67 and Exhibit 3A.

simulation using the sub-regression method finds fully one-third of the pass-through rates to be *negative*, an economically nonsensical result.

By way of contrast, a regression that pools these products into groups results in correct answers. Even if we pool as few as 15 of these 1000 products together in a single pooled regression, then repeat this same experiment 1000 times, all of the estimated pass-through coefficients are positive, and 89% are statistically significant (we reject the hypothesis that they could be as low as zero).

Varying the number of time periods and cost changes per product used in these simulations, similar patterns of results are obtained. Slicing and dicing a dataset into ever tinier groups of products, the sub-regression method will deliver estimates of pass-through rates that are predictably unreliable, with large fractions of statistically insignificant and even negative estimates for a parameter that we know with certainty to equal 100%. A method that pools products together, by way of contrast, delivers consistent, valid, reliable estimates of the true underlying pass-through rate, even with modest numbers of products pooled together.

In conclusion, the sub-regression approach, when inappropriately applied to a product or group of products in a dataset that is insufficiently large, will deliver results that are neither statistically valid nor reliable. Pooled estimates, by contrast, are highly precise and reliable in recovering the known parameter values from the same data.