UNITED STATES DISTRICT COURT
EASTERN DISTRICT OF LOUISIANA

MDL NO. 2328

IN RE: POOL PRODUCTS DISTRIBUTION MARKET ANTITRUST LITIGATION

SECTION: R(2)

JUDGE VANCE
MAG. JUDGE WILKSON

THIS DOCUMENT RELATES TO ALL CASES

ORDER AND REASONS

Plaintiffs move to exclude portions of two reports and related testimony of Dr. John H. Johnson, IV, the proposed expert of defendants Pool Corporation, SCP Distributors LLC, and Superior Pool Products (collectively, "Pool").[1] Specifically, plaintiffs move to exclude as methodologically unreliable the portions of Dr. Johnson's "Rebuttal Expert Report" and "Written Critique of the Supplemental Report of Dr. Gordon Rausser" in which Dr. Johnson "purport[s] to apply the regression model of direct purchaser plaintiffs' expert Dr. Gordon C. Rausser."[2] Dr. Johnson's analysis is sufficiently grounded in economic theory and methods, as well as

[1] R. Doc. 521.

[2] R. Doc. 521-1 at 6.

the facts of this case, to satisfy reliability standards. The Court denies the motion.

I. BACKGROUND

This is an antitrust case that direct-purchaser plaintiffs (DPPs) and indirect-purchaser plaintiffs (IPPs) filed against Pool and the Manufacturer Defendants—Hayward Industries, Inc., Pentair Water Pool and Spa, Inc., and Zodiac Pool Systems, Inc. Pool is the country's largest distributor of products used for the construction and maintenance of swimming pools (Pool Products). The Manufacturer Defendants are the three largest manufacturers of Pool Products in the United States. As defined in DPPs' Second Consolidated Amended Class Action Complaint and IPPs' Third Amended Class Action Complaint, Pool Products are the equipment, products, parts, and materials used for the construction, renovation, maintenance, repair, and service of residential and commercial swimming pools. Pool Products include pumps, filters, covers, drains, fittings, rails, diving boards, and chemicals, among other goods. Pool buys Pool Products from manufacturers, including the three the Manufacturer Defendants, and in turn sells them to DPPs, which include pool builders, pool retail stores, and pool service and repair companies (collectively referred to as

"Dealers").[3] IPPs are pool owners who indirectly purchased Pool Products manufactured by the Manufacturer Defendants and distributed by Pool.[4]

DPPs filed two consolidated amended complaints–the first on June 29, 2012[5] and the second on June 21, 2013[6]–each of which defendants moved to dismiss. Following the Court's orders on those motions, DPPs were left with the following five claims: (1) a Section 1 claim under the *per se* rule involving a horizontal agreement between the Manufacturer Defendants and Pool to fix free freight minimums; (2) three Section 1 claims under the rule of reason involving three separate vertical conspiracies (one between Pool and each Manufacturer Defendant) to exclude Pool's competitors; and (3) a Section 2 attempted monopolization claim against Pool.[7] Pool has moved for summary judgment on these five claims. In addition, DPPs have moved for class certification. On January 27, 2016, the Court granted summary judgment on DPPs' Section 1 horizontal conspiracy claim for lack of evidence of horizontal

---

3 R. Doc. 284 at ¶ 31.

4 *See* R. Doc. 290.

5 R. Doc. 107.

6 R. Doc. 284.

7 *See generally* R. Doc. 250.

collusion.[8] The Court also granted summary judgment on DPPs' three vertical conspiracy claims.[9]

IPPs filed three amended complaints, the most recent on July 16, 2013.[10] Following the Court's order on defendants' motion to dismiss IPPs' state-law claims, IPPs were left with the following claims: California Unfair Competition Law and rule of reason Cartwright Act claims involving three vertical conspiracies (one between Pool and each Manufacturer Defendant); Arizona Antitrust Act claims involving three vertical conspiracies and a claim of attempted monopolization against Pool; Florida Deceptive and Unfair Trade Practices Act claims involving three vertical conspiracies and a claim of attempted monopolization against Pool; and Missouri Merchandising Practice Acts claims based on defendants' alleged anticompetitive agreements to exclude Pool's rivals and Pool's attempted monopolization.[11] On April 29, 2016, the Court granted Pool's motions for summary judgment on IPPs' vertical conspiracy claims.[12]

---

8 R. Doc. 700.

9 R. Doc. 718.

10 R. Doc. 290.

11 R. Doc. 250 at 21- 30, 35.

12 R. Doc. 718.

II. LEGAL STANDARD

When expert testimony offered by one party is subject to a *Daubert* challenge, the Court must act as a "gatekeeper" under Federal Rule of Evidence 702.

A district court has considerable discretion to admit or exclude expert testimony under Rule 702. *See Gen. Elec. Co. v. Joiner*, 522 U.S. 136, 138-39 (1997); *Seatrax, Inc. v. Sonbeck Int'l, Inc.*, 200 F.3d 358, 371 (5th Cir. 2000). Rule 702, which governs the admissibility of expert witness testimony, provides:

> A witness who is qualified as an expert by knowledge, skill, experience, training, or education may testify in the form of an opinion or otherwise if: (a) the expert's scientific, technical, or other specialized knowledge will help the trier of fact to understand the evidence or to determine a fact in issue; (b) the testimony is based on sufficient facts or data; (c) the testimony is the product of reliable principles and methods; and (d) the expert has reliably applied the principles and methods to the facts of the case.

Fed. R. Evid. 702.

In *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, the Supreme Court held that Rule 702 requires the district court to act as a gatekeeper to ensure that "any and all scientific testimony or evidence admitted is not only relevant, but reliable." 509 U.S. at 589; *see also Kumho Tire Co., Ltd. v. Carmichael*, 526 U.S. 137, 147 (1999) (clarifying that the *Daubert*

gatekeeping function applies to all forms of expert testimony). The Court's gatekeeping function thus involves a two-part inquiry into reliability and relevance.

First, the Court must determine whether the proffered expert testimony is reliable. The party offering the testimony bears the burden of establishing its reliability by a preponderance of the evidence. *See Moore v. Ashland Chem. Inc.*, 151 F.3d 269, 276 (5th Cir. 1998). The reliability inquiry requires the Court to assess whether the reasoning or methodology underlying the expert's testimony is valid. *See Daubert*, 509 U.S. at 592-93. The aim is to exclude expert testimony based merely on subjective belief or unsupported speculation. *See id.* at 590.

The Court in *Daubert* articulated a flexible, non-exhaustive, five-factor test to assess the reliability of an expert's methodology: (1) whether the expert's theory can be or has been tested; (2) whether the theory has been subject to peer review and publication; (3) the known or potential rate of error of a technique or theory when applied; (4) the existence and maintenance of standards and controls; and (5) the degree to which the technique or theory has been generally accepted in the scientific community. *Id.* at 593-95. The Supreme Court has emphasized, however, that these factors "do not constitute a 'definitive checklist or test.'" *Kumho*, 526 U.S. at

150 (quoting *Daubert*, 509 U.S. at 593). Rather, district courts "must have considerable leeway in deciding in a particular case how to go about determining whether particular expert testimony is reliable." *Id.* at 152. Courts have also considered whether experts are "proposing to testify about matters growing naturally and directly out of research they have conducted independent of the litigation, or whether they have developed their opinions expressly for purposes of testifying," *Daubert v. Merrell Down Pharms., Inc.*, 43 F.3d 1311, 1317 (9th Cir. 1995), whether the expert has adequately accounted for obvious alternative explanations, *see Claar v. Burlington N.R.R.*, 29 F.3d 499 (9th Cir. 1994), and whether the expert "is being as careful as he would be in his regular professional work outside his paid litigation consulting," *Sheehan v. Daily Racing Form, Inc.*, 104 F.3d 940, 942 (7th Cir. 1997).

A district court's gatekeeper function does not replace the traditional adversary system or the role of the jury within this system. *See Daubert*, 509 U.S. at 596. As the Supreme Court noted in *Daubert*: "Vigorous cross-examination, presentation of contrary evidence, and careful instruction on the burden of proof are the traditional and appropriate means of attacking shaky but admissible evidence." *Id.* The Fifth Circuit has held that, in determining the admissibility of expert testimony, district courts must

accord proper deference to "the jury's role as the proper arbiter of disputes between conflicting opinions. As a general rule, questions relating to the bases and sources of an expert's opinion affect the weight to be assigned that opinion rather than its admissibility and should be left for the jury's consideration." *United States v. 14.38 Acres of Land, More or Less Situated in Leflore Cty., Miss.*, 80 F.3d 1074, 1077 (5th Cir. 1996) (quoting *Viterbo v. Dow Chem. Co.*, 826 F.2d 420, 422 (5th Cir. 1987)).

Nonetheless, expert testimony "must be reliable at each and every step or else it is inadmissible. The reliability analysis applies to all aspects of an expert's testimony: the methodology, the facts underlying the expert's opinion, the link between the facts and the conclusion, *et alia*." *Knight v. Kirby Inland Marine Inc.*, 482 F.3d 347, 355 (5th Cir. 2007) (citation omitted). "Where the expert's opinion is based on insufficient information, the analysis is unreliable." *Paz v. Brush Engineered Materials, Inc.*, 555 F.3d 383, 388 (5th Cir. 2009).

In *Joiner*, the Supreme Court explained that "nothing in either *Daubert* or the Federal Rules of Evidence requires a district court to admit opinion evidence that is connected to existing data only by the *ipse dixit* of the expert." 522 U.S. at 146. Rather, "[a] court may conclude that there is simply too great an analytical gap between the data and the opinion

proffered." *Id.*; *see also LeBlanc v. Chevron USA, Inc.*, 396 F. App'x 94, 98 (5th Cir. 2010).

Second, the Court must determine whether the expert's reasoning or methodology is relevant. The question here is whether the reasoning or methodology "fits" the facts of the case and will thereby assist the trier of fact to understand the evidence. *See Daubert*, 509 U.S. at 591. "[F]undamentally unsupported" opinions "offer[] no expert assistance to the jury" and should be excluded. *Guile v. United States*, 422 F.3d 221, 227 (5th Cir. 2005) (citing *Viterbo*, 826 F.2d at 422).

III. DISCUSSION

Plaintiffs challenge four aspects of Dr. Johnson's opinion, all of which pertain to his conclusion that Dr. Rausser's Common Factors regression and Overcharge regression are fundamentally flawed. The Court begins by summarizing the challenged portions of Dr. Johnson's opinion and methodology. The Court then assesses plaintiffs' challenges to Dr. Johnson's opinion.[13]

---

[13] The Court summarized the relevant law in its earlier order on defendants' challenge to the admissibility of Dr. Rausser's opinion. *See* R. Doc. 701 at 18-26. The challenged portions of Dr. Johnson's opinion mainly address class certification issues—specifically, whether plaintiffs, relying on

A. Dr. Johnson's Opinion and Methodology

1. Dr. Rausser's Opinion

To summarize the challenged portions of Dr. Johnson's opinion and methodology, the Court begins with an overview of Dr. Rausser's opinion, specifically his evaluation of the effect of Pool's pricing on plaintiffs.

Dr. Rausser opined that Pool could charge inflated prices for Pool Products as a result of a mixture of conduct by Pool and the Manufacturer Defendants. In his report, Dr. Rausser conducts a "liability analysis," detailing several instances of conduct by defendants, which in Dr. Rausser's opinion, had an anticompetitive impact on prices in the relevant market and thereby on plaintiffs. Dr. Rausser summarized the alleged effect of defendants' conduct as follows:

> Through its actions, PoolCorp, in collaboration with the Manufacturing Defendants, directly limited competition from rival distributors and reduced the extent to which Dealers (including proposed Class members) could escape PoolCorp's actions. PoolCorp also signaled its ability and commitment to aggressively dissuade entry and expansion. It created a credible threat that enabled it to elevate and maintain its prices across the United States, causing adverse impact across the proposed Class.[14]

Dr. Rausser's testimony, can satisfy commonality and predominance under Rule 23.

[14] Rausser Initial Report, April 10, 2014, at 88.

According to Dr. Rausser, his "liability analysis" (or aggregate causation analysis, as the Court has called it) demonstrates the impact of the defendants' conduct on Pool's prices.

Having concluded that an impact exists, Dr. Rausser then evaluated whether that impact would be "common" across the class. Dr. Rausser concluded that the allegedly supracompetitive pricing "caus[ed] an impact across all Class members,"[15] "regardless of region or product category."[16] To support his opinion, Dr. Rausser conducted a "Common Factors" regression, which sought to identify the key variables that predict variation in Pool Products prices. For example, Dr. Rausser's Common Factors regression includes variables to represent costs, demand, seasonal variation, customer size, and product characteristics. Dr. Rausser contends that because his Common Factors regression can predict 99% of the variation in prices, it shows that "common factors predominate" in determining pricing across the class.[17]

---

15 *Id.* at 84.

16 *Id.*

17 *Id.* at 98-99.

Next, Dr. Rausser used a "before-and-after" analysis to quantify the alleged impact on the class. The before-and-after analysis seeks to quantify damages by comparing two time periods: a "before" time period, when the allegedly anticompetitive conduct was still underway, and an "after" time period, also referred to as the "benchmark" period, when most of the allegedly anticompetitive conduct had supposedly stopped. To compare the two time periods, Dr. Rausser ran his "Common Factors" regression, but added one additional variable, a "dummy" variable, which is set to 1 during the class period and to 0 during the benchmark period. Dr. Rausser calls this regression his "Overcharge" regression. He used November 21, 2011, as the dividing date between the before period (January 1, 2007, to November 20, 2011) and the benchmark period (November 21, 2011, to September 30, 2012), because that is the date the Federal Trade Commission "publicly announced its investigation into Pool Products distribution" and "[d]uring that benchmark period, Pool[] was effectively prohibited from engaging in their previous anticompetitive conduct."[18] A positive and statistically significant coefficient on the overcharge dummy variable would indicate an overcharge attributable to the difference in competitive conditions between

---

[18] *Id.* at 11.

the class period and the benchmark period. Applying his Overcharge regression, Dr. Rausser concluded that "before November 21, 2011, PoolCorp's prices were significantly elevated due to Defendants' anti-competitive conduct."[19]

2. Dr. Johnson's Opinion

In his Rebuttal Expert Report and again in his Written Critique, Dr. Johnson opines that Dr. Rausser's Common Factors Regression (what Dr. Johnson calls the "Impact regression model") "fails to account for customer-specific characteristics."[20] According to Dr. J0hnson, Dr. Rausser's Common Factors regression is flawed because Dr. Rausser "pool[ed] together the sales transactions of all PoolCorp customers—across all geographies and regardless of what products they purchased" to produce an average result across all customers.[21] Dr. Johnson contends that Dr. Rausser's use of averages "does not allow for the possibility that prices paid by *some* direct purchasers may not have responded in the same manner to

---

19 *Id.* at 99.

20 Johnson Rebuttal Expert Report, June 11, 2014, at 110.

21 *Id.* at 112 ¶176.

changes" in the explanatory variables, such as a customer's geographic location.[22]

To support his conclusion, Dr. Johnson applies Dr. Rausser's Common Factors regression to various subsets of the proposed direct purchaser class. Dr. Johnson conducted a regression for each of Pool's geographic divisions, product departments, and customer types (*i.e.*, pool builders, pool retailers, pool service entities, and others).[23] Dr. Johnson's results show that common factors do not predominate in determining pricing across the class. According to Dr. Johnson, individualized factors—geography, product type, and customer type—result in "substantial differences" in the prices paid by the members of the direct purchaser class.[24]

Plaintiffs argue that Dr. Johnson's application of Dr. Rausser's regression models to subsets of the data, rather than the full data set that Dr. Rausser used, is impermissible "data mining." According to plaintiffs, Dr. Johnson has no economic theory for testing the regressions in this manner

---

22 *Id.*

23 *See id.* at 113.

24 *Id.* at 3 ¶ 7.

and that therefore the Court must exclude Dr. Johnson's "separate regressions" as unreliable.[25]

In applying Dr. Rausser's regression models, Dr. Johnson also utilized the available data differently from Dr. Rausser in two ways. First, Dr. Johnson had to address the "unit-of-measure issue." The Court has previously explained the issue as follows. For some sales transactions, Pool records a quantity of "one," indicating that one unit of a product was sold. For others, a quantity of "one" indicates "one case," meaning that twelve (or some other number greater than one) units of the product were sold. Thus, without the unit-of-measure variable, it is impossible to know whether the price given for a particular transaction is the price for one unit or for multiple units.

It is undisputed that unit-of-measure information is not consistently available throughout the relevant time period—specifically, unit-of-measure information is unavailable in Pool's data from June 2012 to September 2012. To address this problem, Dr. Johnson used only those transactions containing unit-of-measure information. In other words, Dr. Johnson

---

[25] R. Doc. 521-1 at 13.

excluded from his regression analysis all sales transactions after May 2012. In a footnote, Dr. Johnson explains that "[t]o rule out that the different time period is driving the difference in the estimated overcharge," Dr. Johnson estimated an additional regression using the full data and obtained similar results.[26] Dr. Rausser conducted a similar analysis in his Supplemental Report, which the Court permitted him to submit to address what effect, if any, the unit-of-measure issue had on his original opinion. Dr. Rausser calculated his Overcharge regression using the totality of the data, as well as the limited, unit-of-measure-specific data "to confirm that th[e] missing information is not somehow driving the results."[27]

Plaintiffs complain that Dr. Johnson's regressions using the unit-of-measure field should be excluded as unreliable. Plaintiffs argue that Dr. improperly "thr[ew] away millions of transactions" when the data did not contain unit-of-measure information, thus eliminating all transactions after May 2012, or approximately forty percent of Dr. Rausser's benchmark period in the Overcharge regression.[28] According to plaintiffs, Dr. Johnson should have either excluded the unit-of-measure field for all transactions in his

---

[26] Johnson Rebuttal Expert Report, June 11, 2014, at 142-43 n. 385, 387.

[27] Rausser Supplemental Report, July 6, 2014, at 8.

[28] R. Doc. 521-1 at 17.

regression analyses or imputed a value for the unit-of-measure where that information was unavailable.[29]

Second, in applying Dr. Rausser's regression models, Dr. Johnson used Pool's "replacement costs" (or "procurement costs"), meaning the price Pool paid for its products from the manufacturers, as an explanatory variable. Dr. Rausser used a proxy for costs by averaging the producer price indices (PPIs) for certain raw materials identified in Pentair's Forms 10-K for each year of the class period.[30] In his Supplemental Report, Dr. Rausser explained that he deliberately avoided using actual cost data in his regression models because "using this procurement cost variable as a cost control, when it is itself contaminated by the alleged acts that the model is intended to investigate, is unreliable."[31] In Dr. Johnson's Rebuttal Report, Dr. Johnson explains that he used actual costs "associated with [acquiring] the particular products in question" because the Dr. Rausser's "proxy variables . . . are not specific to [Pool's] business or the business of the manufacturers of the products at issue."[32] In addition, Dr. Johnson contends that averaging the

---

29 *Id.*

30 Rausser Initial Report, April 10, 2014, at 95-96.

31 Rausser Supplemental Report, July 6, 2014, at 7 n.19.

32 Johnson Rebuttal Report, June 11, 2014, at 107.

four PPIs yields a variable with no economic meaning because the "base periods" applicable to the PPIs vary widely, with one index for example, based to a point in time nineteen years earlier than another.[33]

Plaintiffs argue that Dr. Johnson's use of actual costs as an explanatory variable in a regression that attempts to explain resulting prices is improper due to "endogeneity" or "simultaneity."[34] "Simultaneity occurs when an explanatory variable is determined jointly with the dependent variable in a model." ABA Section of Antitrust Law, Econometrics: Legal, Practical, and Technical Issues 414 (2d ed. 2014) (hereinafter "Econometrics"). "Endogeneity" occurs when an explanatory variable is correlated with an error term in a model. "Endogeneity may exist if the model has . . . simultaneity." *Id.* at 398-99. Plaintiffs contend that the prices manufacturers charge to Pool (*i.e.*, Pool's costs) and the prices Pool charges to its own customers are both determined by the demand for the products from the end users; thus, in plaintiffs' view, endogeneity or simultaneity renders Dr. Johnson's models unreliable because "prices at the manufacturer and distributor levels are jointly determined."[35]

---

33 Johnson Written Critique, October 14, 2014, at 13-14.

34 R. Doc. 521-1 at 17.

35 *Id.*

Dr. Johnson also conducted "statistical tests" of Dr. Rausser's regression models to determine whether any variables omitted in Dr. Rausser's regressions contribute to changes in prices, the dependent variable tested by the regressions.[36] In a series of footnotes, Dr. Johnson explains that these "statistical tests" are "F-test[s] for joint significance" of the variables representing Pool's geographic divisions, product categories, and customer type.[37] Dr. Johnson applied the F-tests to the dataset from Dr. Rausser's workpapers and to the dataset Dr. Johnson prepared to correct "Dr. Rausser's numerous errors," including the unit-of-measure information.[38] According to Dr. Johnson, his statistical testing indicates that Dr. Rausser "inappropriate[ly] pooled together the sales transactions of direct purchasers for smaller subsets of geographies, products, or customer types."[39]

---

[36] Johnson Rebuttal Report, June 11, 2014, at 110 ¶ 174.

[37] *See, e.g.*, Johnson Rebuttal Report, June 11, 2014, at 110 n.331, 146 n.396-397; Johnson Written Critique, October 14, 2014, at 25 n.91, 29 n.97.

[38] *See* Johnson Rebuttal Report, June 11, 2014, at 110 n.331, 146 n.396-397.

[39] *See, e.g.*, *id.* at 9, 83, 92, 110.

Plaintiffs seek to exclude Dr. Johnson's F-tests as "meaningless" because, according to plaintiffs, any statistical significance revealed by the testing results is due only to the large sample size of transactions tested.[40] According to plaintiffs, because Dr. Johnson added several variables to Dr. Rausser's regression models, resulting in millions of transactions analyzed by the F-test, "it was a nearly foregone conclusion" that the results would show that these variables (geography, type of product, and type of customer) had a statistically significant effect on price.[41]

For all these reasons, plaintiffs argue, Dr. Johnson's opinion and related testimony should be excluded.

B. Analysis

The Court addresses each of plaintiffs' four challenges in turn.

1. Sub-regressions for Geographic Division, Product Category, and Customer Type

Dr. Johnson opines that Dr. Rausser's Common Factors regression does not support Dr. Rausser's conclusion of common impact across the class because, when applied to various subsets of the available data, the regression instead reveals that individualized factors (the type of customer, the product

40 R. Doc. 521-1 at 19.

41 *Id.* at 11.

the customer purchased from Pool, and the customer's geographic location) translate to price differences for different members of the proposed class. Plaintiffs argue that Dr. Johnson's regression analyses lack any underlying economic theory, and are therefore unreliable, because Dr. Johnson does not tether the data subsets tested to his definition of the relevant product or geographic market in this case. According to plaintiffs, Dr. Johnson's "arbitrary" selection of data subsets is impermissible "data mining." "Data mining" is defined in various ways, but plaintiffs use the term to mean arbitrarily manipulating available data to achieve a desired result. *See, e.g.*, A.H. Studenmund, Using Econometics: A Practical Guide 184 (6th ed. 2011) ("[D]ata mining involves estimating a variety of alternative specifications *before* [a] 'best' equation has been chosen."); Jeffrey M. Wooldridge, Introductory Econometrics: A Modern Approach 677 (4th ed. 2009) (explaining "data mining" as "try[ing] different models, different estimation techniques, or perhaps different subsets of data" after obtaining "puzzling results" from a "very careful" data model "until the results correspond more closely to what was expected"); Peter Kennedy, A Guide to Econometrics 394 (5th ed. 2003) ("The undesirable version of data mining occurs when one tailors one's specification to the data, resulting in a specification that is

misleading because it embodies the peculiarities of the particular data at hand.").

Pool defends Dr. Johnson's approach as a "sensitivity analysis" of Dr. Rausser's model to determine whether Dr. Rausser's results hold true upon further evaluation. A "sensitivity analysis" involves "estimating a variety of alternative specifications after a potential 'best' equation has been identified" or otherwise "analyzing data in different ways to see how results depend on methods or assumptions." Studenmund, *supra*, at 184; *accord* David H. Kaye & David A. Freedman, Reference Guide on Statistics, Reference Manual on Scientific Evidence 296 (Federal Judicial Center, 3d ed. 2011). Pool contends that Dr. Johnson selected the particular data subsets in an economically rational manner—that is, Dr. Johnson applied Dr. Rausser's regression model to Pool's different geographic areas, product categories, and customer types because "copious documentary and testimonial evidence . . . show[] that economic conditions varied in relation to such differences."[42]

Econometrics literature and antitrust case law support Dr. Johnson's use of sub-regressions to test the results of Dr. Rausser's regression models. For example, an ABA Section of Antitrust Law monograph on econometrics

---

42 R. Doc. 575-2 at 5.

encourages analysts to further test regression results when using regression analyses as a method of proof for classwide impact. According to this source:

> Other statistical tools, including additional regression specifications, may be used to test whether the average effect represented by a single coeffecient from a classwide regression masks widely varying individual effects that require individualized inquiry, or whether it truly reflects common impact.
>
> One such approach is to divide the proposed class into narrowly defined subgroups and construct a series of regressions to test the stability of any estimate of average impact. . . . Further estimating the average effect of the alleged conspiracy separately for different products, geographies, time periods, suppliers, or purchaser types can yield additional insights into classwide impact and the existence and stability of a common method of proof. Estimates of average impact across groups of customers that are similar in magnitude and/or sign may suggest that the alleged misconduct resulted in common impact for all (or nearly all) members of the proposed class. Alternatively, estimated effects that vary widely or are nonsensical would suggest that the alleged misconduct did not result in common impact for all members of the proposed class.

Econometrics, *supra*, at 357-58. Similarly, Professor A.H. Studenmund writes:

> Sensitivity analysis consists of purposely running a number of alternative specifications to determine whether particular results are *robust* (not statistical flukes). Researchers who use sensitivity analysis run (and report on) a number of different reasonable specifications and tend to discount a result that appears significant in some specifications and insignificant in others. Indeed, the whole purpose of sensitivity analysis is to gain confidence that a particular result is significant in a variety of alternative specifications, functional forms, variable definitions, and/or subsets of the data.

Studenmund, *supra*, at 184.

Courts have also recognized the usefulness of sub-regressions in antitrust class actions. Indeed, in *In re Processed Egg Products Antitrust Litigation*, the United States District Court for the Eastern District of Pennsylvania addressed a *Daubert* challenge to Dr. Rausser's Common Factors regression based on the defense expert's "us[ing] Dr. Rausser's regression model on various subsets of the data and f[inding] inconsistent results." 81 F. Supp. 3d 412, 433 (E.D. Pa. 2015). As plaintiffs do here, the plaintiffs in *Processed Egg* argued that the defense expert's sub-regressions constituted "inappropriate data mining." *Id.* The court ultimately found Dr. Rausser's Common Factors regression admissible, but added:

> [J]ust because the Court has found Dr. Rausser's regression model reliable enough for *Daubert* purposes does not mean that Defendants cannot argue that the curious results uncovered by [defendants' expert] make the regression model unconvincing for purposes of class certification, especially if upon more in-depth "mining" on wider sampling, the "curious results" become less a curiosity and more a norm.

Id. at 434.[43] In *In re Air Cargo Shipping Services Antitrust Litigation*, the plaintiffs moved to exclude the opinion of the defendants' expert regarding his "attempts to test the reliability of McClave's [the plaintiffs' expert's] conclusion by running [his] regressions using smaller subsets of the data." No. 06-MD-1175, 2014 WL 7882100, at *11 (E.D.N.Y. Oct. 15, 2014), *adopted by* 2015 WL 5093503 (E.D.N.Y. July 10, 2015). Specifically, the plaintiffs argued that the defendants' expert "ignored the scientific method by running subsets of data with no economic rationale, instead of forming and testing a hypothesis"—an argument identical to plaintiffs' challenges to Dr. Johnson here. *Id.* at 16. In finding the plaintiffs' argument "without merit," the court explained, "[the defendants' expert] was testing McClave's own hypothesis using a finer lens in order to determine whether his results held true across the class. Nothing in Rule 702 precludes a respondent's expert from testing or critiquing the reliability of the movant's expert's work." *Id.*

---

[43] Plaintiffs rely on *Processed Egg* for support for the argument that "Dr. Johnson's contrived subset construction . . . does not meet economic standards and should be excluded as proscribed data mining." R. Doc. 604 at 9. Having explained the facts of *Processed Egg* in context, it is clear that this case does not support excluding Dr. Johnson's approach as unreliable.

In light of the foregoing sources, the Court finds that Dr. Johnson's use of sub-regressions is sufficiently reliable to be admissible under *Daubert*.

The Court also finds that Dr. Johnson's choice of data subsets is sufficiently grounded in the facts of this case to be admissible. As explained, Dr. Johnson conducted regressions across geographic division, product category, and customer type—categories reflected by Pool's sales transaction data. Plaintiffs contend that these classifications are "arbitrary" because Dr. Johnson does not opine that any of these classifications is the relevant product or geographic market in this case, and Dr. Rausser conducted his regressions based on his definition of the relevant market. But as Dr. Johnson explained at his deposition, these subsets are not meant to "direct[ly] overlap with where . . . pricing is determined or how pricing is determined but rather, [Dr. Johnson's regressions] demonstrate[e] that *even at this level of aggregation*, [Dr. Rausser's] model still fails."[44] "There mere fact that the parties' experts disagree on the best way to test the model is no basis for excluding one expert's approach." *In re Air Cargo*, 2014 WL 7882100, at *17. Because Dr. Johnson has "provided explanations for his methodological decisions that appear reasonable and grounded in

[44] Deposition of Dr. John H. Johnson, IV, July 16, 2014, at 260:21-261:3 (emphasis added).

econometrics literature . . . defendants have met their burden of demonstrating that [his] analysis is sufficiently reliable to be admitted." *See In re Vitamin C Antitrust Litig.*, No. 06-MD-1738, 2012 WL 6675117, at *8 (E.D.N.Y. Dec. 21, 2012).

2. Regressions Using the Unit-of-Measure Field

Plaintiffs also seek to exclude the regressions in which Dr. Johnson included sales transactions that contained unit-of-measure information but excluded sales transactions that did not. Plaintiffs argue that Dr. Johnson's approach is "improper" and that Dr. Johnson should have not used the unit-of-measure field at all or should have imputed a value for the unit-of-measure in those transactions where it was missing. Plaintiffs concede that Dr. Johnson validly determined that using the limited data led to similar results as the using the data for the full time period. Nonetheless, plaintiffs argue that Dr. Johnson "got lucky" and that his regressions using limited data should be excluded "regardless of their outcome."[45]

Plaintiffs fail to cite any econometrics literature or case law for their argument that Dr. Johnson's approach is "improper" and unreliable. Indeed, the only authority plaintiffs rely on supports Dr. Johnson's approach. In that

---

45 R. Doc. 604 at 15.

text, the author notes that "complete case analysis," which "simply discard[s]" any observations with missing values, remains "the standard treatment of missing data" and "the most common method in the absence of readily available alternatives." Roderick J.A. Little, Regression with Missing X's: A Review, 87 J. Am Statistical Ass'n 1227, 1229, 1236 (Dec. 1992). Antitrust commentators also note that excluding observations is a reasonable solution. *See* Econometrics, *supra*, at 59 ("One solution to the problem of missing data is to drop observations with missing data."). Moreover, plaintiffs ignore that their own expert determined that the "missing information is not . . . driving the results" of the Overcharge regression. Dr. Johnson's approach to the missing data is therefore not unreliable, and the Court will not exclude his opinion on this basis.

3. Regressions Using Pool's Actual Costs

In applying Dr. Rausser's regression models, Dr. Johnson used Pool's actual costs, as reflected by Pool's internal data, as an explanatory variable to test the dependent variable, price. As noted, plaintiffs complain that these regressions are unreliable due to an "endogeneity" or "simultaneity" problem. In other words, plaintiffs argue that the regressions are unreliable because Pool's actual costs and the prices Pool charges are jointly determined.

Plaintiffs rely on Dr. Rausser's opinion for this conclusion. According to a footnote in his Supplemental Report, Dr. Rausser is concerned about endogeneity for two reasons.[46] First, Dr. Rausser opines that that Pool's prices and the manufacturers' prices are correlated because "the Manufacturing Defendants (and other vendors) enjoyed [increased prices] as a result of restricted competition due to PoolCorp's alleged actions in concert with the Manufacturing Defendants . . . ."[47] Second, Dr. Rausser notes that Pool's prices are correlated with its procurement costs because "PoolCorp applied percentage margins on top of its procurement costs and justified price increases by procurement cost increases."[48] Dr. Rausser asserted at his deposition that "as [procurement costs] go up, they are going to have an influence on those prices, but then there's also a feedback loop from those prices back to the prices that various buyers from the manufacturers are willing to pay."[49] "Feedback" refers to a phenomenon in which "changes in an explanatory variable affect the values of the dependent variable, and changes in the dependent variable also affect the explanatory

---

46 Rausser Supplemental Report, July 6, 2014, at 7 n.19.

47 *Id.*

48 *Id.*

49 *Id.* at 201:20-24.

variable." Daniel L. Rubinfeld, *Reference Guide on Multiple Regression*, Reference Manual on Scientific Evidence 353 (Federal Judicial Center, 3d ed. 2011).

As an initial matter, part of Dr. Rausser's opinion that endogeneity exists here assumes that collusion among Pool, Pentair, Hayward, and Zodiac distorted the prices of the hundreds or thousands of other Pool Products manufacturers. Notably, the Court has dismissed, either on the pleadings or at summary judgment, all claims involving collusion by Pool and the Manufacturer Defendants and among the Manufacturer Defendants *inter se*. Dr. Rausser's second theory explains how Pool's procurement costs affect Pool's prices, but it does not explain how increases in the prices Pool charges to its customers must increase the prices manufacturers charge to all of their customers, including those other than Pool. Thus, while Dr. Rausser refers to a theoretical "feedback loop," he does not point to any record evidence that such feedback, in fact, exists between Pool's prices and Pool's costs.

In Dr. Johnson's opinion, "costs unquestionably affect" the prices Pool charges its customers, "but there is no reverse relationship between the two," and thus endogeneity, simultaneity, and feedback are not issues.[50] In other words, "as a matter of economic logic," if Pool "wanted to raise its prices, it could not (and would not) dictate" that manufacturers also charge Pool higher prices, thereby increasing Pool's costs.[51] Therefore, according to Dr. Johnson, changes in the price Pool charges to its customers do not affect Pool's actual costs. Dr. Johnson, citing record evidence, also disputes Dr. Rausser's contention that Pool applied across-the-board percentage margins to its costs in determining price.[52]

Dr. Johnson's reasoning on this issue is sufficiently reliable to allow the trier of fact to consider his opinion, notwithstanding Dr. Rausser's competing position. As noted, plaintiffs have not provided any empirical evidence for their contention that Dr. Johnson's use of actual costs renders his regression analyses wholly unreliable. And as Dr. Rubinfeld explains in his discussion of simultaneity in the Federal Judicial Center's Manual on Scientific Evidence, "there are no basic direct statistical tests for determining

---

50 Johnson Written Critique, October 14, 2014, at 9-10.

51 *Id.* at 10.

52 *Id.*

the direction of causality; rather, the expert, when asked, should be prepared to defend his or her assumption based on an understanding of the underlying behavior evidence relating to the businesses or individuals involved." *See* Rubinfeld, *supra*, at 322-23. Here, both experts have sufficiently explained the basis for their assumptions about the direction of causality.

The case on which plaintiffs rely to exclude Dr. Johnson's opinion, *In re Polypropylene Carpet Antitrust Litigation*, actually stands for the principle that regressions using a proxy for actual costs and regressions using firm-specific actual costs are *both* admissible. *See* 93 F. Supp. 2d 1348 (N.D. Ga. 2000). In finding the plaintiffs' expert's opinion admissible in *Polypropylene Carpet*, the court rejected the defendants' argument that "firm-specific data necessarily is more appropriate than industry-wide data." *Id.* at 1368. The court also rejected the plaintiffs' argument about the defense expert's approach. The defense expert used internal cost data, including data to which the defendant had "made adjustments." *Id.* The court held that plaintiffs failed to show that the defense expert's use of internal cost data "affect[ed] the reliability of [his] damages estimate." *Id.* The same is true here.

Plaintiffs have not demonstrated that econometricians routinely avoid using actual cost data in conducting regressions to predict changes in price. *See* Fed. R. Evid. 703 (authorizing experts to base an opinion on data upon which "experts in the particular field would reasonably rely"); *Sheehan v. Daily Racing Form, Inc.*, 104 F.3d 940, 942 (7th Cir. 1997) ("*Daubert* . . . requires the district judge to satisfy [herself] that the expert is being as careful as he would be in his regular professional work outside his paid litigation consulting."). Moreover, plaintiffs have not shown that Dr. Johnson's use of actual costs here is unreliable; they have shown only that Dr. Rausser theorizes that it is and that, for this reason, Dr. Rausser chose to use a proxy instead. Because plaintiffs have not pointed to any record evidence or econometrics test that mandates a different conclusion, Dr. Johnson's competing theory about the direction of causality is no basis to find his opinion unreliable. This dispute is a topic for cross-examination.

4. F-Tests

Dr. Johnson relies on F-tests to support his conclusion that Dr. Rausser inappropriately pooled numerous sales transactions, regardless of geographic location, product category, or customer type involved in the transactions, when conducting his Common Factors regression and Overcharge regression. Plaintiffs argue that any statistical significance

shown by Dr. Johnson's F-tests are due only to the large sample size of sales transactions tested. Thus, plaintiffs contend that Dr. Johnson's F-tests are "meaningless" and therefore "irrelevant and unreliable."[53]

Econometrics literature generally encourages the use of F-tests in data analyses. *See, e.g.*, David S. Moore, et al., The Practice of Statistics for Business and Economics 602 (3d ed. 2011) (noting that F statistics test "the null hypothesis that a collection of . . . explanatory variables all have coefficients equal to zero"); Wooldridge, *supra*, at 153 (noting that "[t]esting exclusion restrictions is by far the most important application of F statistics"). For example, antitrust commentators explain that F-tests, or Chow tests,[54] can illustrate whether the effects of an alleged conspiracy should be estimated for certain groups of plaintiffs, rather than collectively estimated across the putative plaintiff class.

> Standard statistical tests can be applied to test the stability of coefficients among subgroups of customers, products, time, geographies, and other subsamples, and to determine whether it is appropriate to pool potential subgroups when estimating the average effect of the alleged conspiracy. For example, a Chow test can be implemented to determine whether the effect of an alleged conspiracy should be estimated separately for two or more potential subgroups of customers, products or periods.

---

53 R. Doc. 521-1 at 19.

54 A Chow test is merely a type of F-test. *See* Wooldridge, *supra*, at 449.

Econometrics, *supra*, 358. The econometrics literature also warns that large sample sizes of data, in and of themselves, may lead to statistically significant results. *See, e.g.*, Kennedy, *supra*, 394 ("Very large sample sizes . . . can give rise to estimated coefficients with very small standard errors [that] may test significantly different from zero, creating a misleading impression of what is important."). What the parties' cited literature does not say, however, is at what point a sample size becomes too large to be helpful in testing the relevant hypothesis.

Here, Dr. Johnson conducted F-tests for subsets of data grounded in the facts of this case. He applied the F-tests using variables representing Pool's geographic divisions, product categories, and customer types, as reflected by Pool's internal sales transaction data. As noted, the F-test is well-accepted in econometrics. Without some evidence that Dr. Johnson used a sample size that is too large—from an econometrics standpoint, not merely as a matter of plaintiffs' assertion—this criticism of his statistical testing results is a cross-examination point. *See generally In re Vitamin C Antitrust Litig.*, No. 06-MD-1738, 2012 WL 6675117, at *8 (E.D.N.Y. Dec. 21, 2012) ("Ask[ing] the Court to take sides in a dispute between experts about the intricacies of econometric modeling . . . is not the proper function of [] *Daubert*."); Manual for Complex Litigation (Fourth) § 23.24 (2004) ("Where

the expert's conclusion is drawn from a reliable methodology . . . the correctness of that conclusion is still an issue for the finder of fact.").

IV. CONCLUSION

In sum, all of plaintiffs' challenges to Dr. Johnson's opinion affect the weight to be assigned to his opinion, rather than its admissibility. *See generally United States v. 14.38 Acres of Land, More or Less Situated in Leflore Cty., Miss.*, 80 F.3d 1074, 1077 (5th Cir. 1996) (quoting *Viterbo v. Dow Chem. Co.*, 826 F.2d 420, 422 (5th Cir. 1987)). Affording the "proper deference to the jury's role as the arbiter of disputes between conflicting opinions," the Court denies plaintiffs' motion to exclude Dr. Johnson's opinion and related testimony. *See id.*

New Orleans, Louisiana, this 12th day of May, 2016.

Sarah Vance

SARAH S. VANCE
UNITED STATES DISTRICT JUDGE