

**UNITED STATES COURT OF INTERNATIONAL TRADE**

**STUPP CORPORATION ET AL.,**

**Plaintiffs and Consolidated  
Plaintiffs,**

**and**

**MAVERICK TUBE CORPORATION,**

**Plaintiff-Intervenor and  
Consolidated Plaintiff-Intervenor,**

**v.**

**UNITED STATES,**

**Defendant,**

**and**

**SEAH STEEL CORPORATION AND  
HYUNDAI STEEL COMPANY,**

**Defendant-Intervenors and  
Consolidated Defendant-  
Intervenors.**

**Before: Claire R. Kelly, Judge**

**Consol. Court No. 15-00334**

**OPINION AND ORDER**

[Sustaining the U.S. Department of Commerce's third remand redetermination in the less-than-fair-value investigation of welded line pipe from the Republic of Korea.]

Dated: February 24, 2023

Jeffrey M. Winton and Jooyoun Jeong, Winton and Chapman PLLC, of Washington, D.C., argued for plaintiff SeAH Steel Corporation.

Robert R. Kieपुरa, Trial Attorney, Commercial Litigation Branch, Civil Division, U.S. Department of Justice, of Washington, D.C., argued for defendant. With him on the brief were Claudia Burke, Assistant Director, Patricia M. McCarthy, Director, and Brian M. Boynton, Principle Deputy Assistant Attorney General. Of Counsel was Mykhaylo Gryzlov, Senior Counsel, Office of the Chief Counsel for Trade Enforcement and Compliance, U.S. Department of Commerce, of Washington, D.C.

Jeffrey D. Gerrish, Schagrin Associates, of Washington, D.C., argued for plaintiff Welspun Tubular LLC USA. With him on the brief were Roger B. Schagrin and Saad Y. Chalchal.

Kelly, Judge: Before the court is the U.S. Department of Commerce's ("Commerce") third remand redetermination in its 2015 less-than-fair-value investigation of welded line pipe imported from the Republic of Korea ("Korea"). See Final Results of Redetermination Purs. Ct. Remand, April 4, 2022, ECF No. 208 ("Remand Results"); see also Welded Line Pipe From [Korea], 80 Fed. Reg. 61,366 (Dep't Commerce Oct. 13, 2015) (final determination of sales at less than fair value), as amended by Welded Line Pipe From [Korea], 80 Fed. Reg. 69,637 (Dep't Commerce Nov. 10, 2015) ("Amended Final Determination") and accompanying Issues & Decisions Memo, A-580-876, (Oct. 5, 2015), ECF No. 30-3 ("Final Decision Memo"). In Stupp Corporation v. United States, the Court of Appeals for the Federal Circuit vacated this court's opinion, remanding to Commerce to further explain why it is reasonable to apply the Cohen's *d* test as part of its differential pricing analysis if certain statistical assumptions have not been met. Stupp Corporation v. United States, 5 F.4<sup>th</sup> 1341 (Fed. Cir. 2021) ("Stupp III"). For the following reasons, the court sustains Commerce's third remand redetermination.

## BACKGROUND

The court presumes familiarity with the facts of this case as set out in this court's previous opinions, as well as the Court of Appeals' decision in Stupp III, and now recounts only the facts relevant to the court's review of the Remand Results. On November 14, 2014, Commerce initiated an antidumping duty investigation of welded line pipe from Korea. Welded Line Pipe From [Korea], 79 Fed. Reg. 68,213, 68,213 (Dep't Commerce Nov. 14, 2014) (initiation of less-than-fair-value investigation). Commerce published its final determination on October 5, 2015 and, finding that 39.72% of SeAH Steel Corporation's ("SeAH") U.S. sales passed the Cohen's *d* test, applied the average-to-transaction method to those sales. Final Decision Memo. at 4. Commerce accordingly calculated a 2.53% dumping margin for SeAH. Amended Final Determination at 69,638. SeAH appealed, arguing that Commerce's differential pricing analysis and application of the Cohen's *d* test were contrary to law and unsupported by substantial evidence. See Stupp Corp. v. United States, 359 F. Supp. 3d 1293, 1302 (Ct. Int'l Tr. 2019) ("Stupp I"), reconsideration denied, 365 F. Supp. 3d 1373 (Ct. Int'l Tr. 2019). SeAH also argued that Commerce improperly rejected its case brief, which contained citations to certain academic texts not part of the administrative record. Id. at 1300–03; Letter from Commerce Rejecting SeAH's Sept. 1, 2015 Case Br., 1–2, PD 384, bar code 3302027-01 (Sept. 3, 2015); [SeAH's] Case Br., PD 377–79, bar codes 3301610-01–03 (Sept. 1, 2015) ("SeAH's Rejected Brief").

This court sustained Commerce's determinations with respect to its use of differential pricing analysis and rejection of SeAH's case brief. Stupp I, 359 F. Supp. 3d at 1299–1306. Specifically, the court found that Commerce correctly rejected SeAH's brief because the academic authorities cited in the brief constituted new factual information intended to advance SeAH's arguments. Id. at 1301. The court also found that Commerce's differential pricing analysis was supported by substantial evidence because, among other reasons, Commerce was not required to apply the Cohen's *d* test in accordance with academic literature. Id. at 1302–06.

The Court of Appeals remanded, instructing Commerce to further explain why its use of the Cohen's *d* test was reasonable in light of "significant concerns" related to application of the test. Stupp III, 5 F. 4<sup>th</sup> at 1357. Specifically, the Court of Appeals questioned the reasonableness of Commerce's application of Cohen's *d* test to data failing to satisfy the statistical criteria of normality, equal variance, and sufficient observation size. Id. 1357–60. Citing to academic literature examining the use of Cohen's *d* test to measure effect size, the Court of Appeals expressed concern that Commerce's failure to satisfy the statistical criteria assumed by Cohen's test could "undermine the usefulness of the interpretive cutoffs," resulting in artificially inflated dumping margins. Id. at 1357. The Court of Appeals affirmed the remaining issues from Stupp I, including this court's decision to uphold Commerce's rejection of SeAH's case brief. Id. at 1344.

## JURISDICTION AND STANDARD OF REVIEW

The court has jurisdiction pursuant to 28 U.S.C. § 1581(c) (2018), which grants the court authority to review actions initiated under 19 U.S.C. § 1516a(a)(2)(B)(i)<sup>1</sup> contesting the final determination in an antidumping duty order. The court will uphold Commerce’s determination unless it is “unsupported by substantial evidence on the record, or otherwise not in accordance with law.” 19 U.S.C. § 1516a(b)(1)(B)(i). “The results of a redetermination pursuant to court remand are also reviewed ‘for compliance with the court’s remand order.’” Xinjiamei Furniture Co. v. United States, 968 F. Supp. 2d 1255, 1259 (Ct. Int’l Tr. 2014).

## DISCUSSION

On remand, SeAH challenges Commerce’s application of the Cohen’s *d* test on the grounds that (1) assumptions underlying the test have not been met, (2) the large cutoff prescribed by the test is arbitrary, and (3) random variables such as exchange rates can cause “false positives.” See Cmts. of [SeAH] on Final Determ. on Remand, 5–36, June 14, 2022, ECF No. 216 (“SeAH’s Cmts.”). Defendant and Welspun Tubular LLC (“Welspun”) counter that (1) the assumptions are inapplicable, (2) Commerce’s application of Cohen’s *d* test leads to reasonable results, (3) the cutoff is supported by statistical literature, (4) SeAH cannot introduce non-record documents for the first time on remand, and (5) SeAH failed to exhaust administrative remedies for its

---

<sup>1</sup> Further citations to the Tariff Act of 1930, as amended, are to the relevant provisions of Title 19 of the U.S. Code, 2018 edition.

exchange rate-related arguments. See Def.'s Corr. Resp. to Cmts., 9–34, Sept. 22, 2022, ECF No. 230 (“Def.’s Reply”); [Welspun’s] Reply [SeAH’s] Cmts. on Remand Redeterm., 18–32, Aug. 15, 2022, ECF No. 218 (“Welspun’s Reply”). For the following reasons, the court sustains the results of Commerce’s remand redetermination.

### **I. SeAH’s Non-Record Documents**

SeAH’s comments to the Remand Results reference several pieces of academic literature which were not included in the administrative record. See SeAH’s Cmts. at 6–36. Welspun and Defendant argue that the court should disregard these materials, as judicial review is limited to the agency record. Welspun’s Reply at 19–20; Def.’s Reply at 10–12. SeAH argues that the court may take judicial notice, or otherwise consider, these materials to better understand the statistical principles behind Cohen’s *d* test. Reply of [SeAH] to Responses by Def. and [Welspun], 10–11, Sept. 28, 2022, ECF No. 236 (“SeAH’s Reply”). For the following reasons, the court need not take judicial notice of SeAH’s non-record documents to understand the statistical principles they illustrate.

Judicial review is generally limited to the administrative record before the agency at the time it rendered its decision. See Camp v. Pitts, 411 U.S. 138, 142 (1973). “The purpose of limiting review to the record actually before the agency is to guard against courts using new evidence to ‘convert the “arbitrary and capricious” standard into effectively de novo review.’” Axiom Res. Mgmt, Inc. v. United States,

564 F.3d 1374, 1380 (Fed. Cir. 2009) (quoting Murakami v. United States, 46 Fed. Cl. 731, 735 (2000), aff'd, 398 F.3d 1342 (Fed. Cir. 2005)).

This court previously upheld Commerce’s decision to reject SeAH’s non-record documents, on the grounds that the submissions constituted new factual information not on the administrative record. Stupp I, 359 F. Supp. 3d at 1299–1306. SeAH’s submissions primarily cited academic articles relating to application of the Cohen’s *d* test under certain conditions. SeAH’s Rejected Brief at 26–33. On appeal, the Court of Appeals affirmed this court’s decision rejecting the non-record information, concluding that SeAH’s materials were not introduced to correct inaccuracies in Commerce’s reporting, but to support its argument challenging Commerce’s use of Cohen’s *d* test. Stupp III, 5 F.4<sup>th</sup> at 1350. In Stupp III, the Court of Appeals nevertheless referenced and quoted from several of the non-record texts introduced by SeAH.<sup>2</sup> Id. at 1357–59. On remand, Commerce asked SeAH to place the

---

<sup>2</sup> The Federal Circuit cited the following five works: Grissom, Robert and Kim, John, Effect Sizes for Research: Univariate and Multivariate Applications (2<sup>nd</sup> ed. 2012), A-580-876, PRRD 8, bar code 4181776-01 (Nov. 12, 2021) (“Grissom & Kim”); Coe, Robert, It’s the Effect Size Stupid: What Effect Size Is and Why It Is Important, paper presented at the Annual Conference of the British Educational Research Association (September 2002), A-580-876, PRRD 8, bar code 4181776-01 (Nov. 12, 2021) (“Coe”); Lane, David, et al., Introduction to Statistics, Online Edition, A-580-876, PRRD 8, bar code 4181776-01 (Nov. 12, 2021) (“Lane”); Algina, James, Keselman, H.J., and Penfield, Randall, An Alternative to Cohen’s Standardized Mean Difference Effect Size: A Robust Parameter and Confidence Interval in the Two Independent Groups Case, 10 *Psychological Methods* (2005), A-580-876, PRRD 8, bar code 4181776-01 (Nov. 12, 2021) (“Algina”); Li, Johnson Ching-Hong, Effect Size Measures in a Two-Independent-Samples Case With Nonnormal and Nonhomogenous Data, *Behavioral Research* (2015), A-580-876, PRRD 8, bar code 4181776-01 (Nov. 12, 2021) (“Li”).

previously-rejected materials on the record, which SeAH did. See Letter from [Commerce] to Interested Parties, A-580-876, PRRD 1, bar code 4176823-01 (Oct. 29, 2021); SeAH Submission of Publications Requested, A-580-876, PRRD 8, bar code 4181776-01 (Nov. 12, 2021). In its comments on the remand redetermination, SeAH again cites new academic sources not on the record, arguing that the court may consider the underlying statistical principles which the texts discuss.<sup>3</sup> See SeAH's Cmts. at 6–25. SeAH claims that the Court of Appeals considered SeAH's previous academic sources in Stupp III, despite upholding Commerce's rejection of SeAH's brief which contained these materials. SeAH's Reply at 10–11. Although SeAH states that the Court of Appeals took judicial notice of the texts, it later clarified that the court may consider the statistical principles regardless of whether the texts themselves are on the record. Response of [SeAH] to Def's Sur-Reply, 2–3, Nov. 14, 2022, ECF No. 247 ("SeAH's Sur-Reply").<sup>4</sup>

---

<sup>3</sup> SeAH cites to the following six non-record sources in its comments: Todd D. Little, Oxford Handbook of Quantitative Methods in Psychology (2013); Ricca and Blaine, Notes on a Nonparametric Estimate of Effect Size, 90:1 *Journal of Experimental Education* 249 (2022); Hedges and Olkin, Overlap Between Treatment and Control Distributions as an Effect Size Measure in Experiments, 21:1 *Psychological Methods* 61 (2016); Huberty and Lohman, Group Overlap as a Basis for Effect Size, 60:4 *Educational and Psychological Measurement* 543 (2000); J. Cohen, A Power Primer, 112:1 *Psychological Bulletin* 155 (1992); F. Alvarez, A. Atkeson, and P. Kehoe, If Exchange Rates Are Random Walks, Then Almost Everything We Say about Monetary Policy is Wrong, Federal Reserve Bank Of Minneapolis Research Department Staff Report 388 (2007).

<sup>4</sup> SeAH argues that the Court of Appeals' decision "stands for the proposition that, when an agency purports to be using a statistical test in accordance with widely-adopted statistical practice, the courts may consider non-record academic materials to evaluate that claim." SeAH's Sur-Reply at 2.



Consistent with the approach of the Court of Appeals, the court may recognize the basic statistical principles discussed in these texts. The idea, for example, that a skewed statistical sample may yield inaccurate results is inductive reasoning—not an assertion of fact. The Court of Appeals’ references to academia do not render its reasoning dependent on academic sources. Thus, the court considers Commerce’s Cohen’s *d* methodology in the same way it would review any other methodology, and may make logical inferences without taking judicial notice of SeAH’s literature.

## II. Administrative Exhaustion

SeAH argues that random fluctuations of exchange rates can affect the Cohen’s *d* test, and lead to inaccurate results. SeAH’s Cmts. at 24–28. Welspun and Defendant argue SeAH failed to properly exhaust this argument. Welspun’s Reply at 20–21; Def.’s Reply. at 27–28. For the following reasons, the court concludes that SeAH has exhausted this argument.

Pursuant to 28 U.S.C. § 2637(d), the court “shall, where appropriate, require the exhaustion of administrative remedies,” including at the preliminary determination stage before the agency. 28 U.S.C. § 2637(d); 19 C.F.R § 351.309(c)(2). Section 2637(d) grants the court “discretion to identify circumstances where exhaustion of administrative remedies does not apply.” ABB, Inc. v. United States, 920 F.3d 811, 818 (Fed. Cir. 2019) (quoting Consol. Bearings Co. v. United States, 348 F.3d 997, 1003 (Fed. Cir. 2003)). The court may also excuse exhaustion in certain

circumstances, such as when a party is raising a “pure question of law.” Agro Dutch Indus. Ltd. v. United States, 508 F.3d 1024, 1029 (Fed. Cir. 2007).

In its comments on the draft redetermination, SeAH argued that random factors, such as exchange rates, could cause the standard deviation of test populations to vary significantly. SeAH’s Cmts. on Draft Redetermination, 17–20, A-580-876, PRRD 30, bar code 4224356-02 (March 21, 2022). SeAH did not provide an exchange rate table, or assert that its actual sales during the period of review were affected by these factors. See id. Subsequently, in its comments on the final remand results, SeAH again argues that Cohen’s *d* could be significantly affected by random factors where the population of data is not normally distributed. SeAH’s Cmts. at 24–28. SeAH adds that its sales were, in fact, affected by fluctuations in the exchange rate between the U.S. dollar and Korean won, because its inland freight expenses were denominated in won. Id. at 24. Welspun counters that SeAH failed to raise its exchange rate argument and supporting factual information during the draft redetermination. Welspun’s Reply at 26.

SeAH has exhausted its exchange rate argument. SeAH’s exchange rate examples provide an illustration of how it believes random factors can render the Cohen’s *d* test inaccurate when values are not normally distributed. Normal distributions is one of the three assumptions that the Court of Appeals remanded to Commerce to explain. See Stupp III, 5 F. 4<sup>th</sup> at 1360. Therefore, Welspun’s argument that Commerce had no opportunity to address SeAH’s exchange rate calculations

misses the point; these calculations are not a new argument, but an illustration of the same scenario Commerce was directed to explain.

SeAH separately argues Commerce must “ignore” fluctuations in exchange rates pursuant to 19 U.S.C. 1677b-1(a). SeAH’s Br. at 8. SeAH concedes that it did not raise this argument in its comments to the draft remand results; nevertheless, it argues that this argument may be considered as a “pure question of law.” Oral Argument at 0:27:39–0:27:53. Defendant argues that Commerce had no opportunity to consider this argument on remand, and Welspun characterizes the argument as a mixed question of law and fact. Oral Argument at 0:24:15–0:26:33, 0:27:56–0:28:39. Whether § 1677b-1(a) is pertinent to Commerce’s differential pricing analysis is a matter of statutory interpretation, not subject to exhaustion requirements. See Agro Dutch Indus., 508 F.3d at 1029. However, SeAH’s argument that § 1677b-1(a) directs Commerce to compensate for exchange rate variations is inapposite. In its full context the statute directs Commerce to use the exchange rate “in effect on the date of sale” for valuation of merchandise, and to ignore fluctuations on that particular date. 19 U.S.C. 1677b-1(a). The plain language does not mandate that Commerce compensate for a respondent’s decision to report expenses in a foreign currency, as SeAH suggests.

### **III. Differential Pricing Analysis**

In Stupp III, the Court of Appeals remanded for further explanation of Commerce’s application of the Cohen’s *d* test as part of its differential pricing

analysis. Stupp III, 5 F.4<sup>th</sup> at 1360. On remand, SeAH renews its argument that Commerce's application of Cohen's *d* test is flawed because it fails to take into account assumptions of sample size, distribution, and variance underlying the test, and further argues Commerce's choice of Cohen's large cutoff is arbitrary. SeAH's Cmts. at 6–24. SeAH also claims random fluctuations in exchange rates can affect the *d* coefficient, causing even test groups with identical prices to pass. Id. at 24–28. Commerce counters that its Cohen's *d* analysis does not operate in a vacuum, and must be considered with the ratio test and meaningful difference test. See Remand Results at 26, 28, 30–31, 41–42, 54–60. Commerce also argues the cutoffs are tied to real-world criteria, that small fluctuations in price will not lead to “false positives” in Cohen's test, and that use of the 0.8 threshold results in reasonably infrequent application of alternative methodologies. Remand Results at 16–19, 32, 54–60. For the following reasons, Commerce has adequately addressed Court of Appeals' concerns.

When investigating whether subject merchandise is being sold at less than fair value, Commerce typically compares “the weighted average of the normal values to the weighted average of the export (and constructed export prices) for comparable merchandise” unless it determines another method is appropriate. 19 U.S.C. § 1677f-1(d)(1)(A)(i); 19 C.F.R. § 351.414(c)(1). This average-to-average (“A-to-A”) method compares the weighted average of a respondent's home country sales prices during the investigation period to the weighted average of the respondent's sales prices in

the United States during the same period. See 19 C.F.R. § 351.414(b)(1). One concern with the A-to-A method is that it may allow a foreign producer or exporter to engage in “targeted dumping,” which occurs when an exporter sells at a dumped price to particular customers or regions, while selling at higher prices to other customers or regions. See Apex Frozen Foods Priv. Ltd. v. United States, 862 F.3d 1337, 1341 (Fed. Cir. 2017) (“Apex II”). As a result, higher-priced products can mask dumped products when Commerce averages the sales using the A-to-A method.

Congress addressed concerns over targeted dumping with the passage of 19 U.S.C. § 1677f-1(d)(1)(B). See Apex II, 862 F.3d at 1342. Section 1677f-1(d)(1)(B) allows Commerce to compare “the weighted average of the normal values to export prices . . . of individual transactions for comparable merchandise if (i) there is a pattern of export prices . . . for comparable merchandise that differ significantly among purchasers, regions or periods of time, and (ii) [Commerce] explains why such differences cannot be taken into account using [the A-to-A method or transaction-to-transaction method<sup>5</sup>].” 19 U.S.C. § 1677f-1(d)(1)(B)(i)–(ii). Targeted dumping is more likely when export prices fit a pricing model that differs significantly across different market segments. Apex II, 862 F.3d at 1341–42. Congress has not provided a method for Commerce to use to determine whether a pattern of significantly different prices

---

<sup>5</sup> Commerce’s regulations provide that the transaction-to-transaction method, which compares prices of individual transactions, will be employed only in rare cases, “such as when there are very few sales of subject merchandise and the merchandise sold in each market is identical or very similar or is custom-made.” 19 C.F.R. § 351.414(c)(2).

exists. However, the Statement of Administrative Action (“SAA”) of the Uruguay Round Agreements Act explains that Commerce should proceed “on a case-by-case basis, because small differences may be significant for one industry or one type of product, but not for another.” Uruguay Round Agreements Act, Statement of Administrative Action, H.R. Doc. No. 103-316, vol. 1, at 842–43 (1994), reprinted in 1994 U.S.C.C.A.N. 4040, 4178.<sup>6</sup>

To determine whether the criteria set forth in § 1677f-1(d)(1)(B) are met, Commerce conducts a “differential pricing analysis” of a respondent’s sales. See Differential Pricing Analysis; Request for Comments, 79 Fed. Reg. 26,720, 26,722 (Dep’t of Commerce May 9, 2014). This analysis contains three tests. First, Commerce applies to respondent’s sales what it refers to as the “Cohen’s *d* test,” described in more detail below, which measures the degree of price disparity between groups of sales. Id. Commerce then counts the percentage of sales by value which “pass” the Cohen’s *d* test, and applies its “ratio test.” Id. at 26,722–23. If 33% of respondent’s sales or less pass, Commerce uses the A-to-A method, and if 66% or more pass, Commerce uses the A-to-T method. Id. If the total percentage of passing sales is between 33% and 66%, Commerce takes a hybrid approach, applying the A-to-T method to those sales passing the test, and the A-to-A method to the remainder. Id. Finally, if Commerce has not selected the A-to-A method for all sales, it applies

---

<sup>6</sup> The SAA is an “authoritative expression by the United States concerning the interpretation and application” of the Uruguay Rounds Agreement Act. 19 U.S.C. § 3512(d).

the “meaningful difference” test to determine whether the A-to-A method could nevertheless account for the disparate pricing. Id. at 26,723. Commerce applies the test by comparing a respondent’s dumping margin using both A-to-A and the selected method. Id. If the A-to-A margin is below the de minimis threshold and the margin from the selected method is not, or if both margins are above the threshold and differ by 25% or more, Commerce continues to use the selected method; otherwise, Commerce applies the A-to-A method for all sales. Id.

As applied by Commerce, the Cohen’s *d* test involves comparing the product-specific prices of “test groups” of a respondent’s sales to a “comparison group” by region, purchaser, and time period. Stupp III, 5 F.4<sup>th</sup> at 1346. For each category, Commerce segregates sales into subsets, with one subset becoming the test group, and the remaining subsets being combined as the comparison group. Id. Commerce then calculates the means and standard deviations of the test and comparison groups. Id. Commerce then calculates a Cohen’s *d* coefficient by dividing the difference in the groups’ means by the groups’ standard deviation.<sup>7</sup> Id. Each subset is thus tested against the remaining subsets across each category, and assigned a *d* coefficient by

---

<sup>7</sup> Thus,  $d = |\text{mean of test group} - \text{mean of control group}| \div \text{standard deviation}$ . Commerce uses a modified version of this formula, substituting the square root of the simple average of the groups’ variances for standard deviation. The Cohen’s *d* test solves for a coefficient representing “effect size.” See generally Cohen, Jacob, Statistical Power Analysis for the Behavioral Sciences, (2<sup>nd</sup> ed. 1988), A-580-876, PRRD 8, bar code 4181776-01 (Nov. 12, 2021) (“Cohen”). “Effect size quantifies the size of the difference between two groups, and may therefore be said to be a true measure of the significance of the difference.” Coe at 7.

solving Cohen's ratio. If the  $d$  value of a test group is equal to or greater than the "large threshold," or 0.8, the observations within that group are said to have "passed" the Cohen's  $d$  test. Id. at 1347.

In Stupp III, the Court of Appeals expressed concern that Commerce's application of Cohen's  $d$  under certain circumstances could undermine the usefulness of the test in less-than-fair-value determinations. Specifically, the Court of Appeals identified three potential scenarios in which use of Cohen's  $d$  could be problematic: first, when the distribution of a respondent's sales data is not normal, second, when the test groups have few data points, and third, when there is minimal variance in a respondent's sales. Stupp III, 5 F.4<sup>th</sup> at 1357–59. The assumption of normality is satisfied when a fixed percentage of the population falls within each standard deviation from the mean—in other words, that a population density graph generally shows a symmetrical, bell-shaped curve. See Starnes, Yates, and Moore, Statistics through Applications, 116 (2005), A-580-876, PRRD 8, bar code 4181776-01 (Nov. 12, 2021). The assumption of size is satisfied when the population is sufficiently large. See Cohen at 21. The assumption of homogeneous variances is satisfied when the standard deviations of test and comparison groups are similar. See Grissom at 68–69. Commerce argued in Stupp III, as it does now, that the three assumptions are only relevant as a matter of statistical significance, and do not apply when analyzing a whole population; the Court of Appeals concluded that this explanation did not fully address its concerns. Stupp III, 5 F.4<sup>th</sup> at 1360.



The Court of Appeals illustrated the problems it identified with the Cohen's *d* test through two hypotheticals. First, the Court of Appeals considered a situation in which Commerce analyzed a group of only eight export sales across four groups, such that each test group would consist of only two sales. Stupp III, 5 F. 4<sup>th</sup> at 1358–59. With groups of such small numbers, the Court of Appeals pointed out that there would be some upward bias in effect size, such that the test would produce more “passing” results, and potentially exaggerate dumping margins. Id. at 1359. The Court of Appeals also observed that a group of only two sales would lack normality. Id. Second, the Court of Appeals described a test group of five sales of about \$100 each, which differed from one-another by up to two cents. Id. Because the standard deviation of such a group would be so small, the Court of Appeals pointed out that the denominator in Cohen's ratio would be drastically reduced, again causing an increase in effect size, and inflating the resulting dumping margin. Id. The Court of Appeals noted that an objective examiner looking at these similar sales prices “would be unlikely to conclude that they embody a ‘pattern’ of prices which ‘differ significantly.’” Id. (citing 19 U.S.C. § 1677f-1(d)(1)(B)(i)).

Commerce reasonably explains that Cohen's *d* test does not operate in a vacuum, but as part of the differential pricing analysis as a whole.<sup>8</sup> Turning first to

---

<sup>8</sup> The parties devote a significant part of their briefings to discussion of (1) the permissibility of using Cohen's *d* test on full populations, and (2) questions of

the assumptions of population size and normalcy, the Court of Appeals questioned whether small sample sizes without normal distributions could “exaggerate” dumping margins by introducing an “upward bias” to effect size. Stupp III, 5 F.4<sup>th</sup> at 1359.<sup>9</sup> Addressing the Court of Appeals’ concerns about population size, Commerce

---

statistical significance versus practical significance. See Remand Results at 11–16, 43–51; SeAH’s Cmts. at 7–10; Def.’s Br. at 12–20; Welspun’s Br. at 21–23; SeAH’s Reply at 17–27; Def’s Sur-Reply at 12–29, 34–35; SeAH’s Sur-Reply at 16–24. Neither question is determinative of whether Commerce’s methodology is reasonable. Both arguments have already been raised before the Court of Appeals, which concluded that they did not resolve its concern over whether the absence of certain assumptions forecloses Commerce’s use of Cohen’s *d* test. Stupp III, 5 F.4<sup>th</sup> at 1360.

Commerce correctly asserts that a “t-test” for statistical significance is used with sampled data, and that Dr. Cohen considered normal distribution and equal variance as necessary assumptions in a t-test. See Cohen at 19; Remand Results at 12–16. However, Commerce improperly reasons that because there is no need for a t-test, there is no basis for the assumptions. Remand Results at 14. Commerce also asserts that SeAH’s assumptions are only relevant as a matter of statistical significance, and that they do not apply because Cohen’s *d* test determines practical significance. Remand Results at 14, 43–45. That these assumptions are required for questions of statistical significance does not answer the question of whether they are also needed to determine practical significance, as the Court of Appeals suggests. See Stupp III, 5 F.4<sup>th</sup> at 1360.

Although SeAH claims that academic sources do not support Commerce’s use of Cohen’s *d* in its differential pricing analysis, this argument is inapposite. SeAH’s decision to substantially advance its arguments using labels taken from statistical literature does not alter the court’s obligation on review. See Soc Trang Seafood Joint Stock Co. v. United States, 321 F. Supp. 3d 1329, 1339 n.13 (2018) (“the fact that Commerce has adopted a methodology based upon a statistical tool known as Cohen’s *d*, and chooses to refer to this methodology as Cohen’s *d*, does not diminish the discretion granted to Commerce”); see also Mid Continent Steel & Wire, Inc. v. United States, 31 F.4<sup>th</sup> 1367 (Fed. Cir. 2022) (“Commerce’s job is not to follow a statistical test as explained in published literature for its own sake, but to implement the statutory mandate to determine when prices of certain groups “differ significantly”).

<sup>9</sup> Although the parties dispute whether such results are really “false positives,” it is undisputed that in at least some instances, groups with as few as two sales have

(footnote continued)

explains that its Cohen's *d* analysis does not stand alone, and operates together with the ratio test and meaningful difference test. See Remand Results at 26, 28, 30–31, 41–42, 54–60. Thus, even if the Cohen's *d* values of small test groups were less accurate than for large test groups, this difference does not by itself render Commerce's use of Cohen's test unreasonable, because the ratio test and meaningful difference test compensate for inaccuracies. See id. Commerce's differential pricing analysis looks at the frequency and impact of effect size to detect targeted dumping—not the effect size alone. See Cohen at 8; Remand Results at 26–28. As Commerce points out in its remand redetermination, the “sole purpose of the Cohen's *d* test” is to determine whether prices “differ significantly” across region, time period, or customer. Remand Results at 41. The “pattern” of export prices which Commerce must find under 19 U.S.C. § 1677f-1(d)(1)(B)(i) is then determined by the ratio test. Id. at 42. The ratio test has already been approved by the Court of Appeals, which found that Commerce's choice of the 33% and 66% thresholds was a “reasonable choice.” Stupp III, 5 F.4<sup>th</sup> at 1355. SeAH's attacks on Cohen's *d* test presuppose that what SeAH claims are “false positives” automatically affect the accuracy of

---

passed Cohen's test. See Stupp III, 5 F.4<sup>th</sup> at 1357; SeAH's Cmts. at 17–19; see Remand Results at 55, 58–59. Identifying results as “false” positives begs the question of what is a false positive. See Remand Results at 59 (“To label this result a ‘false-positive’ does not render the variances inaccurate or erroneous”). SeAH illustrates this situation using data from its own sales, showing how a group of only two sales to a single customer passed Cohen's test, despite SeAH's observation that a visual comparison of the groups on a graph showed those sales to be near the average price. SeAH's Cmts. at 18. Commerce counters that a visual inspection may be inadequate in situations involving complex calculations. Remand Results at 59.

Commerce's differential pricing analysis, when in fact Commerce has allowed for 33% positives before there is any potential effect on a respondent's dumping margins.

Commerce also addresses the Court of Appeals' concern whether samples without normal distributions will produce an inappropriate number of passes. SeAH points to numerous academic sources which it claims confirm the usefulness of Cohen's test is compromised when comparing data sets with non-normal distributions. See SeAH's Cmts. at 7, n.19 (citing Cohen at 13); Id. at 12 (citing Ellis at 41); Id. at 13 (citing Starnes, Yates, and Moore at 135). The Court of Appeals has acknowledged some of these sources. See Stupp III, 5 F.4<sup>th</sup> at 1357–59 (citing Cohen at 21, Grissom & Kim at 66, Coe at 13, Lane at 645, Algina at 318, and Li at 1571). The court need not opine on the relevance of these academic observations;<sup>10</sup> however, it logically follows that a relatively large-tailed distribution (i.e., with large standard deviation) in a test group would tend to decrease Cohen's *d* coefficient, while the opposite would result in an increase. See Remand Results at 29 (“in other words, the fat-tailed distribution may undervalue the significance of effect”) (emphasis in

---

<sup>10</sup> The task of the court is not to interpret the meaning of literature treating with correct application of Cohen's *d*. Rather, the court must determine whether Commerce's methodology is reasonable in light of considerations that run counter to its decision. See Motor Vehicle Mfrs. Ass'n of U.S. v. State Farm Mut. Auto. Ins. Co., 463 U.S. 29, 43 (1983); Ceramica Regiomontana, S.A. v. United States, 636 F. Supp. 961, 966 (Ct. Int'l Tr. 1986), aff'd, 810 F.2d 1137, 1139 (Fed. Cir. 1987) (“As long as the agency's methodology and procedures are reasonable means of effectuating the statutory purpose, and there is substantial evidence in the record supporting the agency's conclusions, the court will not impose its own views as to the sufficiency of the agency's investigation or question the agency's methodology”).

original). SeAH focuses on the second of these two scenarios, arguing that even inputting random data, such as exchange rates, can cause test groups to frequently pass Cohen's *d* test. SeAH's Cmts. at 25–28. SeAH further argues that the ratio test does not account for such random fluctuations.<sup>11</sup> Oral Argument at 0:42:57–0:43:43. Commerce addresses these arguments by explaining that even if Cohen's test can produce positive results under unusual circumstances, this possibility does not mean its use of Cohen's *d* is unreasonable when combined with the ratio test and meaningful difference test. See Remand Results at 26, 28, 30–31, 41–42, 54–60.

The Court of Appeals also specifically asked Commerce to explain why it can use the 0.8 threshold identified by Dr. Cohen as a measure of a significant price difference, when Commerce evaluates data which fails to meet statistical assumptions of normality, size and variance.<sup>12</sup> Stupp III, 5 F.4<sup>th</sup> at 1360. Although Commerce reiterates those assumptions are irrelevant, see Remand Results at 11–

---

<sup>11</sup> Commerce explains that, even if exchange rate fluctuations do affect prices, this effect is not “random” because a respondent can control in which currency it denominates its prices. Remand Results at 45; Oral Argument at 0:49:11–0:50:05.

<sup>12</sup> Although the Court of Appeals approved the 0.8 cutoff in Mid Continent Steel & Wire, Inc. v. United States, it explained in Stupp III that it had yet to consider the reasonableness of the 0.8 cutoff value when the assumptions in question have not been met. Stupp III, 5 F.4<sup>th</sup> at 1356–57 (“We held that . . . it is reasonable to adopt that [0.8] measure where there is no better objective measure of effect size. We did not, however, address SeAH’s second argument [on assumptions] in Mid Continent”) (citation omitted) discussing Mid Continent Steel & Wire, Inc. v. United States, 940 F.3d 662, 673 (Fed. Cir. 2019); see also Oral Argument at 1:39:45–1:40:30. More specifically, SeAH argues that it is unreasonable to compare its prices, which are not normally distributed, using a subjective benchmark that was derived from a normally-distributed population. SeAH’s Cmts. at 10–12.

16, it also explains its choice of the 0.8 threshold as a function of its differential pricing analysis. First, Commerce explains that it employs the 0.8 threshold to identify where prices “differ significantly” pursuant to 19 U.S.C. § 1677f-1(d)(1)(B)(i). Remand Results at 11. Second, Commerce states the 0.8 measurement “represents a difference which is ‘grossly perceptible.’” Remand Results at 52. Reasonably discernible from this statement is that Commerce considers a significant difference to be grossly perceptible in the same way that Dr. Cohen identified a large threshold as one that is “grossly perceptible.” See Cohen at 27. The SAA to the Uruguay Round Agreements Act directs Commerce to proceed “on a case-by-case basis, because small differences may be significant for one industry or one type of product, but not for another,” SAA at 842–43; thus, Commerce’s choice of a measurement that is a function of standard deviation as a uniform approach to identify differences as significant is reasonable, even if the absolute difference in means is small. Commerce’s approach tailors the question of what is a significant difference in price to the pricing parameters of different products. Third, Commerce adequately explains its adoption of Cohen’s widely-recognized choice of 0.8 as a large threshold as significant. Remand Results at 18. It explains that it chose the 0.8 standard because it was “a conservative standard to determine that the observed price differences are significant.” Id. Commerce summarizes its reasoning by explaining that “[u]sing Dr. Cohen’s thresholds is a reasonable approach to interpret whether the difference in the prices is significant and the further interpretation of the

difference in the prices in the context of the calculation of dumping margins ensures the reasonable and limited application of the alternative comparison methodology.” Id. at 33. Thus, Commerce chose a threshold it predicted would result in limited application of the alternative methodology.

Although Commerce adopted this yardstick from Dr. Cohen, and did so because it was widely acknowledged in the statistical literature, Commerce does not rely on the prominence of this yardstick alone. Commerce elaborates that its “actual application of the Cohen’s *d* test in the context of the differential pricing analysis resulted in the application of an alternative comparison methodology to a relatively small number of respondents.” Remand Results at 32. Discernible from Commerce’s explanation is that the 0.8 cutoff produces reasonable passing rates once the ratio and meaningful difference tests are applied. SeAH challenges Commerce’s reliance on the 0.8 threshold as large, arguing that Commerce’s only basis for using the threshold is that it is widely accepted. SeAH’s Cmts. at 10–11. However, in addition to relying on a widely-accepted standard for “grossly perceptible” to determine what is significant, Commerce defines “significant” with reference to the impact a price difference has on a respondent’s dumping margins. Remand Results at 32. Finding that the 0.8 threshold leads to relatively few determinations of targeted dumping, Commerce concludes that its choice is reasonable. Id.

Congress delegated to Commerce the authority to determine where a price difference is significant. 19 U.S.C. § 1677f-1(d)(1)(B)(i). Congress also made clear

that the definition of a “significant price difference” would depend on the product at issue. See SAA at 842–43. Thus, Congress entrusted Commerce to use its expertise and knowledge of pricing to gauge price distinctions. Cf. Fujitsu General Ltd. v. United States, 88 F.3d 1034, 1039 (Fed. Cir. 1996) (granting Commerce significant deference in determinations “involv[ing] complex economic and accounting decisions of a technical nature”). Commerce’s decision to adopt Cohen’s 0.8 (“large”) threshold as a measure of significance because it is widely accepted in the statistical literature does not undermine the reasonableness of that choice, if it is based on Commerce’s expertise and Commerce demonstrates the reasonableness of that choice with reference to the impact it has on the differential pricing analysis. Thus, Commerce’s reference to Cohen’s work does not circumscribe its discretion to choose the same values in a new context, because that choice is itself reasonable.

Commerce addresses the Court of Appeals’ concern that prices with small variances, which hover around the same value, will produce inaccurate results on Cohen’s test. As an initial matter, Commerce explains that results which pass Cohen’s test under these circumstances are not “false positives,” as small differences in average prices will mean that variances, too, will be small. Remand Results at 59; Oral Argument at 1:06:32–1:07:02. Thus, it is discernable that a small variance means a small difference in price will be more significant, and a passing result under these circumstances is not necessarily “erroneous.” Remand Results at 59. Nevertheless, the Court of Appeals observed that an objective examiner considering



a group of sales where prices differed by only a few cents would be unlikely to conclude that they show a “pattern” of prices that “differ significantly” under the statute. Stupp III, 5 F.4<sup>th</sup> at 1359. Commerce responds to this issue by pointing out that an examiner would indeed conclude that there was no pattern—because Commerce does not look for a pattern at this stage of its differential pricing analysis. Remand Results at 41. Again, Commerce explains that the ratio test determines whether a pattern exists, while Cohen’s *d* test only shows whether there are significant price differences. Id. at 41–42. Thus, Cohen’s test would need to generate enough “false positives” to overcome the 33% threshold, at minimum, and there is no evidence on the record suggesting that price patterns, such as that proposed by the Court of Appeals, occur with frequency in SeAH’s sales.

Additionally, to specifically address the hypothetical proposed by the Court of Appeals, Commerce explains that, in addition to the ratio test, the meaningful difference test would prevent low-variance sales which pass Cohen’s *d* test from impacting a respondent’s dumping margins. See Remand Results at 30–31. Adopting the Court of Appeals’ example in which all of a respondent’s prices hovered around \$100 and passed Cohen’s test, Commerce explains that even in this extreme scenario, the respondent would still be assessed under the A-to-A method. Id. Choosing a normal value for comparison equal to the highest sales price, and thus maximizing the respondent’s theoretical dumping margin, Commerce observes the margin would still be under the 2% de minimis threshold. Id.; see 19 U.S.C § 1673d(a)(4),

1673b(b)(3). SeAH argues that Commerce’s reliance on the meaningful difference test is misplaced, because even changes of less than 2% in a respondent’s dumping margin can cross the de minimis threshold and result in a “meaningful difference” finding. SeAH’s Sur-Reply at 15. Specifically, SeAH argues that when the Cohen’s *d* results from small-variance data sets of different products are cumulated, Commerce may find that a respondent’s sales pass the thresholds for both ratio test and the meaningful difference test, even if price differences are negligible. Oral Argument at 1:13:37–1:16:06.<sup>13</sup> This argument overstates Commerce’s position. It is reasonably discernable that Commerce does not rely on the meaningful difference test to prevent all “inappropriate” passes from affecting a respondent’s dumping margins. Commerce has explained the meaningful difference test compensates for a specific concern with low-variance sales which the Court of Appeals identified. See Stupp III, 5 F.4<sup>th</sup> at 1359; Remand Results at 30–31. Moreover, SeAH’s argument is misplaced, because the question before the court is not whether it is possible to construct an unusual scenario where Cohen’s *d* test can result in an alternative comparison method. Rather, the question is whether Commerce’s use of Cohen’s test, when applied as a component of its differential pricing analysis, is reasonable. See Ceramica Regiomontana, S.A. v. United States, 636 F. Supp. 961, 966 (Ct. Int’l Tr.

---

<sup>13</sup> SeAH does not argue that it received an alternative method because its own combined sales inappropriately passed the Cohen’s *d* test. Rather it offers a hypothetical to challenge the reasonableness of Commerce’s methodology more generally. SeAH’s Cmts. at 8, 14–21.

1986). Thus, for the forgoing reasons, Commerce has adequately explained how its methodology is reasonable.

### CONCLUSION

For the foregoing reasons, Commerce's remand results are supported by substantial evidence and comply with the court's Order, Oct. 8, 2021, ECF No. 192, in conformity with the Court of Appeals' Mandate, Oct. 8, 2021, ECF No. 191, and are therefore sustained. Judgment will enter accordingly.

/s/ Claire R. Kelly  
Claire R. Kelly, Judge

Dated: February 24, 2023  
New York, New York