

UNITED STATES COURT OF INTERNATIONAL TRADE

MID CONTINENT STEEL & WIRE, INC.,

**Plaintiff and Consolidated
Defendant-Intervenor,**

v.

UNITED STATES,

Defendant,

and

PT ENTERPRISE, INC. ET AL.,

**Defendant-Intervenors and
Consolidated Plaintiffs.**

Before: Claire R. Kelly, Judge

Consol. Court No. 15-00213

OPINION AND ORDER

[Remanding the U.S. Department of Commerce's third remand redetermination in its antidumping duty investigation of certain steel nails from Taiwan.]

Dated: April 3, 2023

Adam H. Gordon, Jennifer M. Smith, and Lauren Fraid, The Bristol Group PLLC of Washington, D.C., for plaintiff and consolidated defendant-intervenor Mid Continent Steel & Wire, Inc.

Ned H. Marshak, Andrew T. Schutz, and Max F. Schutzman, Grunfeld Desiderio Lebowitz Silverman & Klestadt LLP of Washington, D.C., and New York, N.Y., for consolidated plaintiffs and defendant-intervenors PT Enterprise, Inc., Pro-Team Coil Nail Enterprise Inc., Unicatch Industrial Co., Ltd., WTA International Co., Ltd., Zon Mon Co., Ltd., Hor Liang Industrial Corp., President Industrial Inc., and Liang Chyuan Industrial Co., Ltd.

Mikki Cottet, Senior Trial Counsel, Commercial Litigation Branch, Civil Division, U.S. Department of Justice, of Washington, D.C., for defendant. Also on the brief were Brian M. Boynton, Principal Deputy Assistant Attorney General, and Patricia M. McCarthy, Director. Of counsel was Vania Wang, Attorney, Office of the Chief Counsel for Trade Enforcement & Compliance, U.S. Department of Commerce, of Washington, D.C.

Kelly, Judge: Before the court is the U.S. Department of Commerce's ("Commerce") third remand redetermination in the antidumping duty investigation of certain steel nails from Taiwan, in accordance with the mandate of the Court of Appeals for the Federal Circuit in Mid Continent Steel & Wire, Inc. v. United States, 31 F.4th 1367 (Fed. Cir. 2022) rev'g in part 945 F. Supp. 3d 1298 (Ct. Int'l Tr. 2021). The Court of Appeals vacated and remanded for Commerce to reconsider or further explain its use of a simple average as the denominator of the Cohen's *d* test, as part of Commerce's differential pricing analysis. See Mandate, June 13, 2022, ECF No. 177; Remand Order, June 14, 2022, ECF No. 178. On remand, Commerce again asserts that its use of simple averaging is supported by statistical literature. See Final Results of Redetermination Purs. Ct. Remand, Nov. 10, 2022, ECF No. 186-1. For the following reasons, the court remands Commerce's third remand redetermination for further explanation or reconsideration.

BACKGROUND

The court presumes familiarity with the facts of this case from this court's previous opinions, as well as the Court of Appeals' decision in Mid Continent V, and now recounts only the facts relevant to the court's review of the Remand Results. On June 25, 2014, Commerce initiated an antidumping duty investigation of certain steel

nails from six countries, including Taiwan. See Certain Steel Nails from India, the Republic of Korea, Malaysia, the Sultanate of Oman, Taiwan, the Republic of Turkey, and the Socialist Republic of Vietnam, 79 Fed. Reg. 36,019 (Dep't Commerce June 25, 2014) (initiation of less-than-fair-value investigations). On May 20, 2015, Commerce issued its final determination, which resulted in an antidumping duty order on subject nails from Taiwan. See Certain Steel Nails from Taiwan, 80 Fed. Reg. 28,959 (Dep't Commerce May 20, 2015) (final determination of sales at less than fair value) ("Final Results") and accompanying Issues and Decision Memorandum, May 13, 2015, ECF No. 17 ("Final Decision Memo.").

On March 23, 2017, this court sustained Commerce's determination, including its decision to use a simple average in the denominator of Cohen's *d* test. See Mid Continent Steel & Wire, Inc. v. United States, 219 F. Supp. 3d 1161 (Ct. Int'l Tr. 2017) ("Mid Continent I"). On October 3, 2019, the Court of Appeals vacated this court's judgment and remanded in part to Commerce for further explanation of its decision to use the simple average in Cohen's *d* test. See Mid Continent Steel & Wire, Inc. v. United States, 940 F.3d 662 (Fed. Cir. 2019) ("Mid Continent III"). On remand, Commerce defended its decision to use the simple average, explaining that its use of the simple average was both accurate and in accord with statistical literature. See Final Results of Redetermination Purs. Ct. Remand, June 16, 2020, ECF No. 144-1 ("Second Remand Results"). On January 8, 2021, this court again sustained, concluding that Commerce had adequately explained how its use of simple averaging

was more accurate, and thus a reasonable choice of methodology. See Mid Continent Steel & Wire, Inc. v. United States, 945 F. Supp. 3d 1298 (Ct. Int'l Tr. 2021) (“Mid Continent IV”). On April 21, 2022, the Court of Appeals again vacated this court’s judgment, remanding to Commerce for further explanation of its decision to use the simple average. See Mid Continent Steel & Wire, Inc. v. United States, 31 F.4th 1367 (Fed. Cir. 2019) (“Mid Continent V”).

On remand, Commerce again defends its decision to use the simple average with the Cohen’s *d* test, explaining that its usage is consistent with statistical literature. See Final Results of Redetermination Purs. Ct. Remand, Nov. 10, 2022, ECF No. 186-1 (“Remand Results”). Consolidated Plaintiffs and Defendant-intervenors PT Enterprise, Inc., et al. (“PT”) submitted comments asserting that Commerce’s use of the simple average is not supported by literature and resulted in increased dumping margins, as well as challenging Commerce’s decision to exclude certain of its submissions from the record. See [PT’s] Cmts. on Remand Results, Dec. 13, 2022, ECF No. 188 (“PT’s Cmts.”). Plaintiff and Consolidated Defendant-intervenor Mid Continent Steel & Wire, Inc. (“Mid Continent”) submitted comments supporting Commerce’s use of simple averaging. See [Mid Continent’s] Cmts. Supp. Remand Results, Feb. 13, 2023, ECF No. 191 (“Mid Continent’s Cmts.”). On February 27, 2023, PT moved for oral argument, see [PT’s] Mot. Oral Arg., Feb. 27, 2023, ECF

No. 198, and on March 21, 2023, Mid Continent submitted its response in opposition to oral argument. See Resp. Opp. Oral Arg., March 21, 2023, ECF No. 200.¹

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)² 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

PT’s Rejected Submissions

As a preliminary matter, PT argues that Commerce improperly rejected portions of its case brief as new factual information. PT’s Cmts. at 31. Commerce rejected a report authored by a statistical consultant for PT, which was submitted together with PT’s comments on the draft remand results. See Rejection Ltr., ECF No. 195, A-583-854, PRRD 15, bar code 4304452-01 (Oct. 25, 2022) (“Rejection Ltr.”);

¹ In light of the court’s decision on the merits, PT’s motion for oral argument will be denied as moot.

² 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.

see also W.A. Huber Decl. Concerning Draft Results of Redetermination Pursuant to Court Remand, ECF No. 195, A-583-854, PRRD 11, bar code 4290765-02 (Sept. 30, 2022) (“Huber Decl.”).

Commerce properly rejected the Huber Declaration. Commerce’s regulations require that it reject untimely-filed factual information unless the record is reopened or, where appropriate, an extension is sought. 19 C.F.R. § 351.302(d). “Factual information” includes, among other things, data or statements of fact in support of allegations. 19 C.F.R. § 351.102(b)(21)(ii). Expert reports submitted to Commerce “analyzing reported information clearly assume[] the weight of evidence and, as such, amount[] to [d]ata or statements of fact in support of allegations, i.e., factual information.” PSC VSMPO-Avisma Corp. v. United States, 688 F.3d 751, 760 (Fed. Cir. 2012) (internal quotation and citation omitted). Commerce did not reopen the record for new submissions on remand, see Defendant’s Resp. Cmts. on Remand Redetermination, 29–30, Feb. 13, 2023, ECF No. 192 (“Def.’s Reply”), and the Huber Declaration contains both the analysis and conclusions of PT’s statistical expert as to the validity of the remand results. See generally Huber Decl. Therefore, the report constitutes new factual information, which was properly rejected by Commerce.

PT also challenges Commerce’s rejection of certain pages it submitted from Cohen, and references in its comments to two non-record academic sources. PT’s Cmts. at 32–33; Rejection Ltr. at 3. PT argues that these sources are not factual information, but are only references to information in the public realm. PT’s Cmts.

at 32–33. Defendant counters that academic literature is factual information which must be submitted as part of the record for Commerce to consider. Def.’s Reply at 31. Again, factual information includes “data or statements of fact in support of allegations,” 19 C.F.R. § 351.102(b)(21)(ii), and PT’s academic sources contain statistical information which PT cites to support its argument in favor of weighted averaging. See PT’s Cmts. at 32. Therefore, as with the Huber declaration, Commerce did not reopen the record for new submissions on remand, so Commerce properly rejected PT’s untimely submissions.

Cohen’s *d* Test and the Simple Average

In Mid Continent V, the Court of Appeals remanded to Commerce to further explain or reconsider its use of the simple average in the Cohen’s *d* test, in light of statistical literature suggesting that only use of a weighted average is appropriate. Mid Continent V, 31 F.4th at 1381. Commerce makes two arguments to support its use of the simple average. First, it argues that the literature supports using the simple average when the “full population” of data is considered. Remand Results at 14. Second, it argues that sample sizes are irrelevant, because they only affect “statistical significance.” Id. at 14–16.³ PT counters that the literature only supports

³ Commerce also argues against the alternative weighting proposals advanced by PT and the Court of Appeals. Remand Results at 16–23. Because Commerce has not elected to use either method on remand, the court does not reach the relative merits of using either a single standard deviation or PT’s proposed equation. Moreover, even if the alternate proposals were fully supported by the literature, this support would not necessarily detract from Commerce’s choice of the simple average. See Motor Vehicle Mfrs. Ass’n of U.S. v. State Farm Mut. Auto. Ins. Co., 463 U.S. 29, 43 (1983).

weighted averaging, and that for Commerce’s purposes, analysis of full populations is no different from analysis of a sample. PT’s Cmts. at 5–18. For the following reasons, the court remands Commerce’s determination.

When investigating whether merchandise is being sold at less than fair value, Commerce typically compares the weighted average of normal values with the weighted average of 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). To address concerns over “targeted dumping,”⁴ Congress amended 19 U.S.C. § 1677f-1 to allow 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 [with another method].” 19 U.S.C. § 1677f-1(d)(1)(B)(i)–(ii). Congress has not specified how Commerce should determine whether a “pattern of significantly different prices” exists.⁵ Therefore, the court determines whether Commerce’s methodology is

⁴ Targeted dumping occurs when an exporter sells at a lower, “dumped” price to particular customers or regions, while selling at higher prices to other customers or regions, such that the higher-priced products mask the dumped products by increasing the overall average price. See Apex Frozen Foods Priv. Ltd. v. United States, 862 F.3d 1337, 1341 (Fed. Cir. 2017).

⁵ The Statement of Administrative Action of the Uruguay Round Agreements Act

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). See also, e.g., Stupp Corporation v. United States, 5 F.4th 1341, 1354 (Fed. Cir. 2021) (standard for reviewing components of Commerce's differential pricing methodology is reasonableness) (citing Mid Continent III, 940 F.3d at 667).

To implement § 1677f-1(d)(1)(B) Commerce performs 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—the Cohen's *d* test, the ratio test, and the meaningful difference test. Only the Cohen's *d* test, which determines whether there is a “pattern of significantly different prices,” is at issue in this case.

As applied by Commerce, the Cohen's *d* test involves comparing the prices of “test groups” of a respondent's sales to a “comparison group” by region, purchaser, and time period. See id. at 26,722. 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

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.

and standard deviations of the test and comparison groups. Id. Commerce finally calculates a d coefficient by dividing the difference in the groups' means by the square root of the average of the squared standard deviations of each group.⁶ See, e.g., Xanthan Gum From the People's Republic of China: Final Determination of Sales at Less Than Fair Value, 78 Fed. Reg. 33,350 (Dep't of Commerce June 4, 2013). Commerce finds the average of the squared standard deviations by adding them together and dividing by two, referring to the result as a "simple average." See id. Commerce does not account for the differences in the size of each group, i.e., use a "weighted average."

Commerce tests each subset against the remaining subsets across each category, and assigns a d coefficient by dividing the difference in the groups' means by the square root of the average of the squared standard deviations of each group. 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. Differential Pricing Analysis; Request for Comments, 79 Fed. Reg. 26,720, 26,722 (Dep't of Commerce May 9, 2014). If a sufficient quantity of sales by volume pass Cohen's d test, Commerce may compare the export prices of individual transactions

⁶ Thus, $d = |m_A - m_B| / \sqrt{(\sigma_A^2 + \sigma_B^2)/2}$, where $|m_A - m_B|$ is the absolute value of the difference in means between the test and comparison groups, and $\sigma_A^2 + \sigma_B^2$ is the sum of the squared standard deviation of both groups. Standard deviation squared (σ^2) is also referred to as "variance." Commerce's formulation of what it calls the Cohen's d test is also known as Cohen's equation (2.3.2). See Cohen, Jacob, *Statistical Power Analysis for the Behavioral Sciences*, 44, (2nd ed. 1988), A-580-876, PRRD 8, bar code 4181776-01 (Nov. 12, 2021) ("Cohen").

to normal value, instead of comparing the average export prices to normal value. Id. at 27,622–23.

In Mid Continent V, the Court of Appeals held that Commerce inadequately explained its choice of the simple average over the weighted average in constructing the Cohens' *d* denominator, given that the literature relied upon by Commerce only supports weighted averaging. Mid Continent V, 31 F.4th at 1378–81. Although Commerce is not ordinarily bound to follow published literature, because it justified its methodology by relying on that literature, the Court held that Commerce needed to justify its departure from established statistical practice.⁷ Id. As the Court of Appeals explained, Commerce's methodology must reasonably implement its statutory mandate of determining when prices of certain groups differ significantly. Id. at 1580–81; see also Motor Vehicle Mfrs. Ass'n, 463 U.S. 29, 43; Ceramica Regiomontana, 636 F. Supp. 961, 966. The Court of Appeals concluded that Commerce's methodology departed from the acknowledged literature. See Mid Continent V, 31 F.4th at 1381.

⁷ Reviewing the literature, the Court of Appeals referenced affirmative language in Cohen, Coe, and Ellis which supports the use of a weighted average. See Mid Continent V, 31 F.4th at 1378; see also Cohen at 67; Ellis, Paul, The Essential Guide to Effect Sizes: Statistical Power, Meta-Analysis, and the Interpretation of Research Results, 26–27, (2010), A-580-876, PRRD 8, bar code 4181776-01 (Nov. 12, 2021) (“Ellis”); Coe, Robert, It's the Effect Size: Stupid: What Effect Size Is and Why It Is Important, 6, (September 2002), A-580-876, PRRD 8, bar code 4181776-01 (Nov. 12, 2021) (“Coe”). The Court of Appeals concluded that Commerce had departed from “all the cited statistical literature governing Cohen's *d*,” and that “[t]he cited literature nowhere suggests simple averaging for unequal-size groups.” Id. at 1377, 1380.

The Court found insufficient Commerce's explanations that (i) the pricing behaviors of the test and comparison groups represented "separate and equally rational" and "equally genuine" behavior, and (ii) using the simple average provides greater "predictability" in antidumping determinations.⁸ *Id.* at 1379–80 (discussing Second Remand Results at 8–9). The Court emphasized Commerce's failure to explain the connection between the rationality and genuineness of a seller's choices and "the undisputed purpose of the denominator figure." *Id.* Relatedly, the Court questioned Commerce's assertion that the means and variances of both test and comparison groups "constitute an abstract effect" which "exclusively define[s] the independent pricing behavior of each group." *Id.* at 1380; Second Remand Results at 45.⁹ The Court held Commerce failed to explain why focusing on the difference

⁸ In context, Commerce explained that:

For the purpose of this particular analysis, Commerce finds that these two distinct pricing behaviors [of the test and comparison groups] are separate and equally rational, and each is manifested in the individual prices within each group. Therefore, each warrants an equal weighting when determining the "standard deviation" used to gauge the significance of the difference in the means of the prices of comparable merchandise between these two groups. Because Commerce finds that each of these pricing behaviors are equally genuine when considering the distinct pricing behaviors between a given purchaser, region, or time period and all other sales, an equal weighting is justified when calculating the "standard deviation" of the Cohen's *d* coefficient.

Second Remand Results at 8–9.

⁹ The Court concluded that the section of Cohen cited by Commerce for the "abstract effect" proposition did not call for simple averaging, and in any case, related to calculating a different measure of effect size, *f*, instead of Cohen's *d*. *Mid Continent V*, 31 F.4th at 1380.

between the groups calls for a simple averaging yardstick. Mid Continent V, 31 F.4th at 1380.

Further, the Court of Appeals concluded Commerce’s assertion that using the simple average provides greater “predictability” was both unclear and inadequate to support its determination. The Court of Appeals stated it was unclear if Commerce was referring to “predictability” as the ability to predict consequences, and if so, it did not see how Commerce had provided a basis for its assertion of greater predictability, given that weighting is calculated on the basis of product weight, value or units, and not number of transactions.¹⁰ Id.

On remand, Commerce abandons its arguments justifying the use of a simple average apart from the literature, and instead argues, again, that the literature supports its methodology. See, e.g., Remand Results at 57 (arguing the choice of a simple average is reasonable because it is supported by the academic literature); id. at 42 (responding to the Court of Appeals’ discussion regarding the qualitative factors and stating “the academic literature does contain support for the use of a simple average”); id. at 42–43 (responding to the Court of Appeals’ direction to explain its choice as reasonable or to reconsider, and explaining that the academic literature supports its choice). Commerce does not further explain how the pricing behaviors of

¹⁰ Additionally, the Court of Appeals had already reached the same conclusion regarding predictability in Mid Continent III, when Commerce claimed that simple averaging was more predicable than weighted averaging. See Mid Continent III, 940 F.3d at 674.

the test and comparison groups represent “separate and equally rational” and “equally genuine” behavior. Commerce likewise does not provide additional support for its assertion of “predictability.”

On remand, Commerce fails to offer any further explanation for its assertions of predictability, abstract effect, and rationality, observing that “[t]he CAFC has already opined in Mid Continent V that these prior arguments are unpersuasive to support that the simple average is reasonable, and now Commerce has taken a new approach which focusses on the academic literature” Remand Results at 36. Commerce misconstrues the Court of Appeals’ mandate. The Court held that Commerce’s non-academic arguments were unsupported—not unsupportable. See, e.g., Mid Continent V, 31 F.4th at 1379 (“Commerce has not offered an adequate explanation of why [equal rationality and genuineness] support[] the particular step Commerce must justify”); id. (“And in any event, Commerce has not provided a reasonable explanation for this predictability assertion”).

Mid Continent attempts to fill the void left by Commerce and argues that there are practical justifications for using the simple average. For example, Mid Continent argues that the “prevalence and sophistication of many respondents’ ‘dump-proofing’ activities” means that using a weighted average could potentially open the door to manipulation of dumping calculations. Mid Continent’s Cmts. at 11. Yet, even if such explanations could support the reasonableness of Commerce’s choice, they are not Commerce’s explanation, and thus they cannot support its determinations. See

SEC v. Chenery Corp., 332 U.S. 194, 196 (1947) (agency action reviewed on grounds invoked by agency).

Instead of attempting to comply with the Court of Appeals' direction "to provide a reasonable justification for departing from what the acknowledge literature teaches," Commerce interprets the Court of Appeals' direction as allowing it to offer an explanation, again, of its view of the literature. Even assuming the Court of Appeals left this option open, Commerce's arguments fail to support its position. First, Commerce argues that the simple average is supported because it uses the full population of sales, and does not estimate means or standard deviations for the test and comparison groups. Remand Results at 14. Therefore, because the literature only contemplates using the weighted average approach when the standard deviations are estimates, Commerce argues that the simple average is supported, and the weighted average is not. Id.

This argument fails for two reasons. First, it is unclear what supports Commerce's premise that Cohen contemplated using equation (2.3.2), the simple average, with full populations.¹¹ Cohen's power tables appear to use sample size as an input, see Cohen at 28, and specify that they are to be used "for power analysis in

¹¹ PT argues that section 2.3 of Cohen's book exclusively concerns sampling from a population. PT's Cmts. at 9. However, as Defendant argues, PT failed to raise this argument in its comments on the preliminary results, so it has not been exhausted. See 28 U.S.C. § 2637; 19 C.F.R. § 351.309(c)(2). Because the court is remanding the matter for further consideration or explanation, Commerce should address this argument in the first instance on remand.

the case where two samples, each of n cases, have been randomly and independently drawn from normal populations.” Id. at 19. Commerce’s reference to equations (2.2.1) and (2.2.2) as “explicitly” calculating effect size based on actual populations seems inconsistent, given that Cohen used these equations to generate d values to create his power tables, not as stand-alone tests. See id. at 20; Remand Results at 35. A test for full populations in the context of power analysis would be redundant on its face, as there would be no question of statistical significance to analyze. Thus, Commerce does not explain, and it is not discernable why Commerce believes that equations (2.2.1) and (2.2.2)—still less equation (2.3.2), which expressly implicates sample size—are intended for testing full populations.¹² Moreover, the Court of Appeals has already held that the literature does not suggest simple averaging for unequal-sized groups. Mid Continent V, 31 F.4th at 1380.

Additionally, even if equation (2.3.2) could be used with full populations, Commerce offers no support for its argument that use of the simple average is reasonable in this context. Rather, Commerce argues that the weighted average would be unreasonable, asserting that the literature discussing weighted averaging is exclusively concerned with sampling. Remand Results at 14; see Cohen at 67.

¹² Defendant argues that PT failed to exhaust its argument that Cohen does not state equation (2.3.2) applies only to full populations. See Def.’s Reply at 15; PT’s Cmts. at 9. In Mid Continent V, the Court of Appeals discusses the use of equation (2.3.2) with sample groups, rather than full populations, implicitly recognizing that the equation does not apply only to full populations. See Mid Continent V, 31 F.4th at 1378. Therefore, because the Court of Appeals has already addressed this issue, the court need not consider whether PT’s argument has been exhausted.

Commerce also asserts that the differential pricing analysis does not involve sampling, but uses full populations, and thus concludes that weighted averaging is inappropriate in light of this distinction. Remand Results at 14. However, Commerce's premise does not lead to its conclusion. That weighted averaging is supported when sampling is present does not mean that it is unsupported when sampling is absent.

Commerce further claims that it may use equation (2.3.2) regardless of sample size, because sample size is only an important factor in the determination of statistical significance. Id. at 15. In support of this proposition, Commerce references Coe, Cohen, and Ellis, who agree that sample size is a necessary input when calculating pooled standard deviation from sampled data. Remand Results at 15; see also Cohen at 40; Coe at 10; Ellis at 10. Commerce's argument again fails for two reasons.

First, although it is true that sample size is necessary to determine statistical significance, it does not follow that sample size is irrelevant where statistical significance is absent. Commerce explains that Cohen's d is one of the three variables (d , a , and n) used to determine statistical significance with a t-test. Remand Results at 16. Thus, Commerce argues, Cohen's warning that power values "may be greatly in error" if both sample sizes and standard deviations are unequal does not apply to the differential pricing analysis, which does not calculate power values. Id. As with its "full population" argument, the fact that Commerce is not conducting a power

analysis does not necessarily mean that it may disregard Cohen's limitations on equation (2.3.2) for calculating d .

Commerce's assertion that equation (2.5.2) requires estimation from a sample while equation (2.3.2) does not require estimation from a sample, appears inconsistent with the literature. See id. at 15 (citing Cohen at 66–67). Although Commerce identifies σ_A and σ_B in equation (2.3.2) as representing standard deviations of full populations, it fails to consider that the σ values themselves seem to be used by Cohen as pre-test estimates of the full population value, which will later be calculated with sampling. See Cohen at 44 (stating that equation (2.3.2) is accurate "provided that sample sizes are about equal"); see also Ellis at 10, 10 n.8 (stating of Cohen's d test "[i]f [the standard deviations] of both groups are roughly the same then it is reasonable to assume that they are estimating a common population standard deviation"). Thus, Commerce's assertion that sampling is not implicated in equation (2.3.2) is unsupported, as Cohen seems to use this equation in calculating statistical power.

Additionally, Commerce's assertion that the literature provides no support for the weighted average appears to contradict Cohen, Ellis, and Coe at a number of points, as the Court of Appeals has already observed. See Mid Continent V, 31 F.4th at 1378 ("the cited literature uniformly teaches use of a pooled standard deviation that involves weighted averaging"). For example, Cohen's power tables use a single n value, representing the size of both samples. See Cohen at 40 ("*Sample size, n.*

This is the size of *each* of the two samples being compared”) (emphasis in original). The tables appear to be designed for what Cohen calls “Case 0,” which is when $\sigma_A = \sigma_B$ and $n_A = n_B$ (i.e., when standard deviations are equal and sample sizes are also equal) See id. at 27. Cohen’s “Case 1” (equation (2.3.1)) and “Case 2” (equation (2.3.2)) represent adaptations to Cohen’s power tables, which allow them to function even when an assumption is not met. See id. at 42–44. As Cohen’s examples following each equation demonstrate, however, all three equations are evidently intended for experimental planning.¹³ See id. None of Cohen’s many illustrative examples show using simple averaging with unequal samples. Equation (2.5.2), on the other hand, is used for calculating d from experimental data. See Cohen at 67 (“Here, our focus shifts from research planning to the appraisal of research results, and . . . the palpable characteristics of the sample”). Ellis and Coe both expressly prescribe equation (2.5.2) for situations where effect size is being calculated from experimental data.¹⁴ See Ellis at 10 n.8; Coe at 10. Neither author discusses using a simple, unweighted average. Therefore, as the Court of Appeals found in Mid Continent V, Commerce’s

¹³ Example 2.3 following equation (2.3.1) shows the power that a psychological experimenter could find in a test, given posited values for α , d , n_A , and n_B . Cohen at 43. Example 2.4 is similar, and shows the power an economist could expect from running an experiment with posited d , α , and n values. Id. at 45. These examples show that the power tables in section 2.3 of Cohen’s book allow an experimenter to find any one value of t , d , α , and n , provided that the other three variables are fixed. See id. at 14. Thus, it appears that equation (2.3.2) is used as a tool to estimate d in order to find one of the other three variables in a proposed experiment.

¹⁴ Coe and Ellis express pooled standard deviation using different notation than Cohen, but their formulae are algebraically equivalent. Compare Ellis at 10 n.8 and Coe at 10 with Cohen at 67.

claim that academia supports the simple average appears to be contradicted by the literature itself. If Commerce continues to rely on the academic literature to support its methodology, it must further explain why its choice of the simple average is reasonable in light of this inconsistency. The matter is remanded to Commerce for further explanation or reconsideration.

CONCLUSION

In accordance with the foregoing, it is

ORDERED that Commerce's determination is remanded for further explanation or reconsideration consistent with this opinion; and it is further

ORDERED that Commerce shall file its remand redetermination with the court within 90 days of this date; and it is further

ORDERED that the parties shall have 30 days to file comments on the remand redetermination; and it is further

ORDERED that the parties shall have 30 days to file their replies to the comments on the remand redetermination; and it is further

ORDERED that the parties shall file the joint appendix within 14 days after the filing of replies to the comments on the remand redetermination; and it is further

ORDERED that Commerce shall file the administrative record within 14 days of the date of filing of its remand redetermination; and it is further

ORDERED that PT's motion for oral argument, see ECF No. 198, is denied as moot.

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

Dated: April 3, 2023
New York, New York