**KUTAK ROCK LLP**

SUITE 1000
1101 CONNECTICUT AVENUE, N.W.
WASHINGTON, D.C. 20036-4374
202-828-2400
FACSIMILE 202-828-2488
www.kutakrock.com

ATLANTA
CHICAGO
DENVER
FAYETTEVILLE
IRVINE
KANSAS CITY
LITTLE ROCK
LOS ANGELES
MINNEAPOLIS
OKLAHOMA CITY
OMAHA
PHILADELPHIA
RICHMOND
SCOTTSDALE
WICHITA

LIZBETH R. LEVINSON
lizbeth.levinson@kutakrock.com
(202) 828-2400

June 23, 2014

**Via Electronic Mail (ECWeb@trade.gov)**

Mr. Paul Piquado
Assistant Secretary for Enforcement and Compliance
U.S. Department of Commerce
Central Record Unit, Room 7045
14th Street & Constitution Avenue, N.W.
Washington, D.C. 20230

**Re: Differential Pricing Analysis; Request for Comments**

Dear Mr. Piquado:

On behalf of the Seafood Exporters Association of India ("SEAI"), we hereby submit comments regarding the use of the differential pricing analysis in response to the Department of Commerce's Federal Register notice of May 9, 2014. *See Differential Pricing Analysis; Request for Comments,* 79 Fed. Reg. 26,720 (Dep't of Commerce May 9, 2014) ("Request for Comments"). These comments are being timely filed no later than June 23, 2014.

In its Request for Comments, the Department of Commerce ("Department") indicated it was seeking comments on its "differential pricing" analysis currently being applied in less-than-fair-value investigations and certain administrative reviews. As the Department notes, "in the first stage of the differential pricing analysis, the Department uses two tests – the "Cohen's *d* test" and the "ratio test" – to determine whether there is a pattern of prices that differ significantly. Our comments will explain why we believe the Department's current methodology is flawed and should not be applied in investigating price differences that might indicate targeted dumping.

## I. SUMMARY OF SEAI'S COMMENTS

A host of legal and statistical considerations demonstrate the Department's differential analysis methodology is seriously flawed. The methodology is contrary to law for the following compelling reasons:

1. The differential pricing methodology is unlawful because the Department did not comply with the notice and comments requirements of the APA codified in 5 U.S.C. § 553(b)(A) before implementing the policy;

2. Under 19 U.S.C. § 166f-(d)(1)B)(ii), the Department lacks the authority to engage in a targeted dumping analysis in administrative reviews as that authority has been expressly confined by Congress to original investigations;

3. As determined by the Court of International Trade in *Gold East Paper (Jiangsu) Co., Ltd.*[1], the Department's failure to follow the notice and comment procedures required by the APA prior to withdrawing the targeted dumping regulation, rendered that withdrawal ineffective. The Department must therefore follow its targeted dumping regulation in current and upcoming proceedings, including only applying the average-transaction comparison method to any sales it may determine are targeted.

[1] 918 F. Supp. 2d 1317 (Ct. Int'l Trade 2013)

4. The Department's current methodology fails to determine whether "export prices… differ significantly" among purchasers, regions, or periods and whether such differences constitutes a pattern. In addition, in its application of the differential pricing analysis to date, the Departments has failed to fulfill its statutory obligation to provide a meaningful explanation of why certain minimal price differences are *significant* under 19 U.S.C.§ 1677 and more specifically failed to evolve any standard to determine how "significant" price differences should be determined.

5. Cohen's d test is inappropriate for determining significant price differences that point to potential targeted dumping for the following reasons:

a. The grouping of regions and purchasers by the entire period of review and of periods by quarters is not in conformity with law and is otherwise not reasonable.

b. The Cohen's d test does not differentiate between allegedly targeted sales and the myriad of other potential causes of variations in price such as normal market price variations.

c. The use of the same set of data in the Test Group and the Base Group creates circular calculations, is not consistent with statistical principles, is a significant deviation from the Department's practice under the *Nails* test, contradicts the Department's own position and is otherwise not reasonable.

d. the Cohen's d test has a tendency to produce "false positives" in the context of targeted dumping;

e. Under Cohen's d, a difference that appears large in terms of the standard deviation may be very small in terms of actual prices and thus is not indicative of *significant* price differentials as required by 19 U.S.C. § 1677-f (1);

f. The Cohen's d test does not distinguish between instances in which the mean prices of the targeted group are above or below the means prices of the base group even though the targeted dumping statute is meant to identify pricing that is aberrationally low.

A full analysis of the economic and statistical flaws associated with the differential pricing analysis is set forth in the memorandum attached hereto as Exhibit A, containing the Verified Statement of Dr. Henry B. McFarland, economic consultant with the firm of Economists, Incorporated. Dr. McFarland's statement is relied upon and cited in support of the legal arguments contained herein.

## II. SEAI'S COMMENTS

### A. The Department Lacks The Authority to Apply the Differential Pricing Analysis in Administrative Reviews

#### 1. The Statute Limits The Alternative Methodology to Investigations

Section 777A of the Tariff Act of 1930 instructs the Department to calculate antidumping margins in investigations by comparing either weighted average prices in the comparison market to weighted average prices in the United States ("A-A") or individual transactions in the comparison market to individual transactions ("T-T") in the United States.[2] Congress prescribed only one exception to this rule in investigations for circumstances in which: (1) there is a pattern of export prices that differ significantly among purchasers, regions, or periods of time; and (2) such difference cannot be taken into account using the standard A-A or T-T methodologies.[3] In such instances, the Department may compare weighted average prices to individual transaction prices ("A-T") to counteract so-called "targeted dumping."

---

[2] 19 C.F.R. §1677f-1(d)(1)(A).

[3] 19 U.S.C. §1677f-1(d)(1)(B).

Section 777A does not provide any parallel provision for addressing targeted dumping in administrative reviews such as here. The statute specifically places the targeted dumping provision under the subheading entitled "Investigations"[4] and not also under the "Reviews" section that immediately follows.[5] By including a specific exception for targeted dumping in investigations, but excluding a similar exception for administrative review, Congress demonstrated its intent to limit the targeted dumping exception to investigations.

The legislative history and overall statutory scheme of the targeted dumping provision are consistent with Congressional intent to limit the application of targeted dumping to investigations only. The Uruguay Round Amendments Act ("URAA"), the implementing statute for the WTO antidumping code, instructed Commerce to compare average prices in the comparison market to average prices in the United States in calculating the duty in investigations. Congress was concerned, however, that this new averaging methodology would lead exporters to mask selective or "targeted" dumping. The Statement of Administrative Action ("SAA") explained this concern:

> In part the reluctance to use an average-to-average methodology had been based on a concern that such a methodology could conceal "targeted dumping." In such situations, an exporter may sell at a dumped price to particular customers or regions, while selling at higher prices to other customers or regions.[6]

---

[4] 19 U.S.C. §1677f-1(d)(1)

[5] 19 U.S.C. § 1677f-1(d)(2)).

[6] Statement of Administrative Action, accompanying the Uruguay Round Agreement Act, House Doc. 103-316 at 656 ("SAA")

The need to create an exception to the average-to-average methodology to counteract targeted dumping was also explained in the SAA:

> New Section 777A(d)(1)(B) provides for a comparison of average normal values to individual export prices . . . in situations where an average-to-average or transaction-to–transaction methodology cannot account for a pattern of prices that differ significantly among purchasers, regions or time periods, *i.e. where targeted dumping may be occurring.*[7]

Targeted dumping was viewed as an issue in investigations only, where the A-A methodology could be manipulated to disguise dumping. There was no mention anywhere in the legislative history of a similar concern for administrative reviews. There is no evidence that a targeted dumping exception for administrative reviews was contemplated or implicitly granted. . . The A-A methodology, which was the root of the fear of targeted dumping, was not used in administrative reviews.

Pursuant to Section 777A(d)(2), the Department previously calculated antidumping margins in reviews by comparing monthly average prices in the comparison market to individual transactions ("MA-T") in the United States, and zeroing all individual negative dumping amounts in the calculation of the final weighted average dumping margin. This methodology would not allow for targeted dumping because individual transactions, rather than averages, were being examined.

The SAA states that: "{t}he Agreement reflects the express intent of the negotiators that the preference for the use of an average-to-average or transaction-to-transaction comparison be limited to the "investigation phase" of an antidumping proceeding."[8] By contrast, the preferred

[7] *Id.*at 843.(emphasis added)
[8] SSA at 843.

methodology in administrative reviews is to compare monthly averages in the United States to individual transactions in the comparison market.[9]

In short, at the time that Section 777A was enacted, there was no need for a targeted dumping exception for administrative reviews. Congress enacted an exception for targeted dumping in investigations because the methodology was A-A, which could mask dumping. The same fear was not present for administrative reviews, where the preferred methodology was MA-T. The plain language of the statute, coupled with the legislative history, evidence a clear Congressional intent to confine the targeted dumping analysis to investigations. Absent legislative authority to apply the targeted dumping exception in administrative reviews, the Department has no authority to do so under the guise of agency discretion.

### 2. There Is No Gap in the Statute for the Department to Fill

The Department's use of the targeting dumping regulations in administrative reviews cannot be justified on the grounds that the statute does not preclude its actions and therefore its interpretation of statutory silence should be accorded deference under *Chevron U.S.A. Inc. v. Natural Resources Defense Council*.[10] First, Congress has not been silent on the issue. Its statutory scheme expressly provided for a targeted dumping analysis as an exception to the normal comparison methodology in investigations, and explicitly chose not to provide for this exception in administrative reviews. The plain language of the statute is unambiguous. Under

---

[9] Id.

[10] 467 U.S. 837 (1984).

generally accepted principles of statutory interpretation "when the legislature uses certain language in one part of the statute and different language in another, the court assumes different meanings were intended." [11] Likewise, where Congress has included specific language in one section of a statute but omitted it from another related section, it should be presumed that Congress intended the omission. [12]

Second, the Court of International Trade has stated that "*Chevron* does not allow an agency to fill gaps in a statute where there is no ambiguity, or take authority upon itself where no such authority has been explicitly or implicitly granted." [13] This principal was upheld in FAG Italia S.p.A. v. United States , where the Federal Circuit distinguished between ambiguous statutory language that creates a "gap" in the statute that an agency may reasonably fill, and a silence in the statute from which an agency cannot create authority. [14] The Federal Circuit noted that no court has held that "an administrative agency has authority to fill gaps in a statute that exist because of the absence of statutory authority."[15]

It is a basic canon of statutory construction that where Congress included specific language in one section of a statute but omitted it from another related section, it should be

[11] Cf. 19 U.S.C. § 1677f-1(d)(1)(B) with 19 U.S.C. § 1677f-1(d)(2). *See Ad Hoc Committee of AZ-NM-TX-FL Producers of Gray Portland Cement v. United States*, 13 F.3d 398, 401-402 (Fed. Cir. 1994)

[12] *See Ad Hoc Committee of AZ-NM-TX-FL Producers of Gray Portland Cement v. United States,* 13 F.3d 398, 401-402 (Fed. Cir. 1994))..

[13] *Marine Harvest v. United States*, 244 F. Supp. 2d 1364, 1379 (Ct. Int'l Trade 2002).
[14] 291 F.3d 809 (Fed. Cir. 2002).

[15] Id. at 815.

presumed that Congress intended the omission.[16] As the Federal Circuit explained in Sioux Honey Association v. Hartford Fire Insurance Company[17]:

> Further, that § 1585 does not contain a "jurisdiction" term is telling, especially because the Customs Courts Act *does* refer to "jurisdiction" numerous times in neighboring provisions (i.e., §§ 1581-1584). "[W]here Congress includes particular language in one section of a statute but omits it in another section of the Act, it is generally presumed that Congress acts intentionally and purposely in the disparate inclusion or exclusion." *Russello v. United States*, **464 U.S. 16**, **23**, **104 S.Ct. 296**, 78 L.Ed.2d 17 (1983) (internal quotation marks omitted). Congress's use of the term "jurisdiction" in §§ 1581-1584 but not in § 1585 suggests that it did not intend for the "powers" term in § 1585 to incorporate the concept of supplemental jurisdiction.

There is no legislative or statutory history that suggests that Congress intended for Commerce to apply targeted dumping to administrative reviews, or that Congress even contemplated such application. To the contrary, the statutory scheme demonstrates that Congress intended for the exception for targeted dumping to apply only in investigations. The legislative history and the SAA further substantiates that the targeted dumping provision was not meant to be applied in administrative reviews because it was neither necessary nor required. Although Commerce has wide discretion in administering the antidumping law, it cannot exercise that discretion contrary to congressional intent.[18]

---

[16] *See Ad Hoc Committee of AZ-NM-TX-FL Producers of Gray Portland Cement v. United States, 13* F.3d 398, 401-402 (Fed. Cir. 1994)

[17] 672 F.3d 1041 (Fed. Cir. 2012)

[18] See *GPX International Tire Corp. v. United States, 666 F.3d 732, 745 (Fed. Cir. 2002)*

### 3. The Department's Failure To Properly Implement the WTO Findings on Zeroing Pursuant to Section 129(b) Has Caused Any Alleged "Gap" With Regard to Targeted dumping

The Department has acknowledged that Section 777(A)(d)(1)(B) does not provide authority to employ the targeted dumping analysis in administrative reviews.[19] The Department nevertheless held that because 19 C.F.R. § 351.414(c)(1) provides that the average-to-average methodology is to be used in both investigations and reviews, that the segments are sufficiently analogous to warrant consideration of targeted dumping in both.

Until recently, the Department had almost always calculated dumping margins in administrative reviews by comparing monthly average prices in the comparison market to individual transactions in the United States, and then zeroing all individual negative dumping amounts in the calculation of the final weighted average dumping margin. A long line of WTO cases subsequently struck down the use of zeroing in both investigations (where the A-A approach is used) and in administrative reviews (where the MA-T method is used).[20] The WTO Appellate Body ruled that the zeroing in antidumping duty reviews (where the MA-T method is used) is inconsistent with Article 9.3 of the Agreement on Implementation of Article VI of the GATT 1994 and Article VI:2 of the WTO General Agreement on Tariffs and Trade 1994, either "as such" or "as applied."

[19] Decision Memorandum for the Preliminary Results of the 2012-2013 Administrative Review of the Antidumping Duty Order on Certain Frozen Warmwater Shrimp from India, at 5.

[20] *See e.g.* US-Zeroing (EC), WT/DS294/R, WT/DS294/AB/R, para. 263(a)(i).

Section 129 of the Uruguay Round Agreement Act ("URAA") governs the Department's implementation of the WTO rulings. Specifically, Section 129(b)(2) provides that "notwithstanding any provision of the Tariff Act of 1930," within 180 days of a written request from the U.S. Trade Representative ("USTR"), the Department shall issue a determination that would render its "action.... not inconsistent with the findings of the Panel or the Appellate Body".[21] Pursuant to Section 129, the United States announced its intention to comply with the WTO adverse decisions on zeroing.

To comply fully, the Department should have abandoned zeroing in administrative reviews using the MA-T method in order to render its action "not inconsistent" with the findings of the appellate body. Instead, the Department abandoned the use of the MA-T methodology altogether in administrative reviews and adopted a new comparison methodology: monthly average to monthly average ("MA-MA").[22] While the Department eliminated zeroing in its new MA-MA calculations in administrative reviews (even though the applicability of zeroing in this context had never been addressed by the WTO), the agency also quietly reintroduced zeroing in investigations where targeted dumping had been identified.

Section 129 does not vest the Department with the authority to determine how to implement the WTO findings. Once a decision is taken to render the Department's actions consistent with the finding of a Panel or Appellate body, the implementation must conform to the judicial findings. Since the Appellate Body found in this instance that zeroing in the MA-T

---

[21] *See* 19 U.S.C. § 3538(b)(2)

[22] *See Antidumping Proceedings: Calculation of the Weighted-Average Dumping Margin and Assessment Rate in Certain Antidumping Proceedings: Final Modification*, 30 Fed. Reg. 8101 (Feb. 14, 2012).

comparison methodology was inconsistent with the Antidumping Duty Agreement, the Department was required to eliminate zeroing in the MA-T comparison methodology. Instead, the Department eliminated zeroing in its new MA-MA calculations in administrative reviews and reintroduced zeroing in the MA-T comparisons in investigations and reviews where targeted dumping had been identified.

The Department claims that investigations and review are analogous because both employed the average-to-average methodology. The Department itself created the analogy by unnecessarily changing the default methodology in reviews from A-T to MA-MA in order to parallel its practice in investigations. In this process, the Department failed to comply with Section 129 under which is purportedly acted.

The Department further claims that a "gap" exists that justifies the use of targeted dumping in reviews. To the extent a gap exists, however, it too is of the Department own making. Had it properly implemented Section 129(b) by eliminating zeroing, but retaining the A-T methodology in administrative reviews, there would be no so-called "gap." Stated simply, the alleged gap upon which the Department relies is a direct result of the agency's flawed implementation of Section 129(b). The Department itself has manipulated the system in a backdoor attempt to introduce targeted dumping and zeroing into administrative reviews. In so doing, it has impermissibly ignored the WTO rulings, the targeted dumping statute, its own regulations and relevant case law.

### B. Even If the Department Has the Authority, Targeting Dumping Is Not Necessary Because Of The Shorter Monthly Averaging Periods In Administrative Reviews

Even assuming that the Department has the authority to apply the targeted dumping methodology in administrative reviews, there is no need for targeted dumping in reviews. The forces that drive targeted dumping in investigation are absent from administrative reviews, despite the Department's insistence that the segments are analogous. The average-to-average methodologies applied in original investigations and administrative reviews are significantly different, as discussed below.

In investigations, the average-to-average methodology is based on an annual average for both normal value and export prices, Under 19 U.S.C. § 166f-1(d)(2) and C.F.R. § 351.414 (d)(3) Commerce must apply shorter monthly averaging periods of time in calculating the margin in administrative reviews. The concern that underlines targeted dumping, namely that a purchaser will take advantage of the annual average methodology to cancel out lower priced sales with higher-priced sales to particular purchaser, regions or times, simply is not an issue when monthly averages are used. The potential to manipulate prices within a short monthly period is significantly less than when an annual period is used.

In other words, the use of the shorter than annual averaging periods in administrative reviews is itself a significant safeguard against targeted dumping as it accounts for significant variations in export prices. Petitioner's acknowledged in the 2010-2011 administrative review of Certain Frozen Warmwater Shrimp from India that the targeted dumping methodology captures at best "period differences."[23] While period differences may be meaningful when an annual average methodology is used, they are simply not so where a shorter period like monthly averages are being examined.

The plain language of 19 C.F.R. §351.414(d)(3) indicates that Congress intended to account for significant price differences not through targeted dumping, but by prescribing shorter comparison periods. The regulation provides:

> When applying the average-to-average method in an investigation, the Secretary normally will calculate weighted averages for the entire period of investigation. However, when normal values, export prices or constructed export prices differ significantly over the course of the period of investigation, the Secretary may calculate weighted averages for such shorter period as the Secretary deems appropriate. When applying the average-to-average method in a review, the Secretary normally will calculate weighted averages on a monthly basis and compare the weighted-average normal value for the contemporaneous month.

This provision thus instructs Commerce to normally calculate the POR-wide averages of prices. An exception is carved out, however, for instances in which "export prices differ significantly." Section 777A (19 U.S C. § 1677f-1(d)(1)(B)(ii)), 19 C.F.R. 351.414(d)(3) and the SAA[24] all use the same phrase "differ significantly."

The prescribed remedy for measuring dumping where export prices differ significantly is to calculate weighted averages for a shorter period. The regulation provides further that the default methodology for calculating margins in reviews is to base both normal value and U.S. price on monthly averages.[25] The administrative review methodology is thus inherently designed

---

[23] See Certain Frozen Warmwater Shrimp from India: Final Results of Antidumping Duty Administrative Review and Final No Shipment Determination, 78 Fed.Reg. 424992 (July 16, 2013)
[24] SSA at 656.

[25] 19 C.F.R. §351.414(d)(3)

to capture the effect of export prices that "differ significantly." Because the existing methodology for reviews already captures targeted dumping, a separate targeted dumping provision is neither necessary nor warranted. Given the safeguards of the shorter comparison period, it is entirely logical that Congress would have concluded that a targeted dumping provision for reviews simply is unnecessary.

### C. <u>If a Targeted Dumping Analysis Is Required in Administrative Reviews, Then It Must Be Developed Through A Formal Rulemaking Process</u>

The APA requires that legislative rules, defined as those that are issued "pursuant to a statutory authority and which implement the statute,"[26] must be promulgated through formal rulemaking procedures.[27] These procedures require an agency to issue a notice of proposed rulemaking with a description of the rule, to solicit, receive and consider comments, and after a full analysis of the legal impact on the regulatory community, to issue the final rule. The Courts broadly construe legislative rules to be those that "create new law, rights or duties in what amounts to a legislative act." [28] A legislative rule does not merely interpret prior regulations; rather, it imposes obligations and has the force of law.[29] If the Department deems that targeted dumping is required in administrative reviews, its analysis must therefore be the subject of a formal rulemaking proceeding. The Department's development of the differential pricing

---

[26] *See the Attorney General's Manual on the Administrative Procedures Act* (1947)

[27] *5 U.S.C. § 553*

[28] *See Sweet v. Sheahan*, 235 F. 3d 80,91 (2nd Cir. 2000)

[29] *American Mining Congress v. Mine & Safety Health Admin, 995 F.2d1106,1109 (D.C. Cir. 1993)*

analysis pursuant to the statutory authority allegedly vested in the agency by 19 U.S.C. § 1677f-1, meets the definition of a legislative rule.

The Department itself has recognized the legal requirement and intrinsic value of a rulemaking by publishing this Request for Comment and also recently publishing proposed and final regulatory provisions not to apply the previously withdrawn regulations governing targeted dumping in investigations.[30] Although we commend the Department's initiative, the opportunity to provide comments on differential pricing should have occurred years ago, prior to having been applied in so many investigations and administrative review.

To date, despite an emphasis generally on transparency, the Department has for the most part developed the targeted dumping methodologies behind closed doors. Despite assurances to the public that the targeted dumping regulations would be promulgated in an open fashion, through rulemaking, interpretative rulings, policy bulletins *etc.*, the Department has kept the issue firmly under wraps.[31] Without explanation, let alone notice and comment, the Department abandoned the Nails test upon which it previously relied. Likewise, the Department developed the differential pricing analysis and began relying on the Cohen's d test without notifying the public or soliciting comments. The proper application of targeted dumping is a critical issue for respondents, and we are pleased that the Department has finally sought comments from the trade community at large.

---

[30] *Non-Application of Previously Withdrawn Regulatory Provisions Governing Targeted Dumping in Antidumping Duty Investigations*, 79 Fed. Reg. 22371 (Apr. 22, 2014).

**D. The Department's Differential Pricing Methodology Does Not Comply With The Regulations, the SSA and Section 777A**

**1. The Department Must Comply With Its Own Targeted Dumping Regulations Which Remain Effective**

Even assuming arguendo that the Department has the authority to engage in a targeted dumping analysis in an administrative review, it must abide by the previous regulations it attempted unlawfully to withdraw in 2008.[32] The Department's withdrawal of its previous targeted regulations failed to comply with the notice and comment requirement of the APA. The withdrawal was therefore ineffective and the regulations remain in effect.

The Department's targeted dumping regulations were previously codified in its regulations at 19 C.F.R. § 351.414(f). The withdrawn regulations provided in part that where the Department found targeted dumping, it would "normally" limit the application of the average-to-transaction method to those sales that constitute targeted dumping." 19 C.F.R. § 351.414(f)(2) ("Limiting Rule). In December 2008, however, the Department suddenly withdrew its targeted dumping regulations without the notice and comment required by the APA.[33] The Department claimed that it was relieved of its obligation to provide notice and comment[34] pursuant to the provisions of 5 U.S.C. § 553(b)(B), which permits an agency to

---

[32] 73 Fed Reg 74930

[33] *See* Withdrawal of the Regulatory Provisions Governing Targeted Dumping in Antidumping Duty Investigations, 73 Fed. Reg. 74,930 (Dec. 10, 2008).

[34] Federal agencies are bound by Section 553 of the APA to issue notice of a proposed rule and consider comments from interested parties in determining its final rule. See 5 U.S.C. § 553. These notice and comment requirements must be followed unless waived "when the agency for good cause finds . . . that notice and public procedure thereon are impracticable, unnecessary, or contrary to the public interest."

dispense with notice and comments where "good cause" exists. The Department further claimed that the good cause exception was applicable because notice and comment was "impracticable and contrary to the public interest." [35] Id. at 74, 931.

The Court of International Trade has held that none of the reasons cited by Commerce in favor of withdrawal rose to the level required for exemption from basic requirements of notice and comment. In *Gold East Paper (Jiangsu Co., Ltd.) v. United States*, 918 F. Supp. 2d 1317 (Ct. Int'l Trade June 17, 2013), the Court held:

> Because Commerce failed to provide notice and comment before withdrawing the Limiting Rule, and the agency failed to provide adequate cause to qualify under the exceptions to the notice and comment requirements, the court finds that the repeal of the regulation was invalid, and the Limiting Rule is still in force. Commerce's decision to apply the targeted dumping remedy to all [respondent's] sales failed to comply with applicable law. Commerce must, on remand, reconsider its application of the targeted dumping remedy under the Limiting Rule. Assuming the finding of targeted dumping remains positive, after reconsideration of the other issues addressed in this opinion, Commerce must limit application of the targeted dumping remedy to the targeted sales, or provide an adequate explanation why the situation is not a "normal" one before applying the remedy to all [respondent's]. sales.

In short, the Department's failure to follow the notice and comment procedures required by the APA prior to withdrawing its targeted dumping regulations nullifies the withdrawal. The

---

These procedures must be followed both when promulgating regulations and when "repealing a rule." 5 U.S.C.§ 551(5).

[35] Section 553(b)(B) of the APA specifically authorizes federal agencies to dispense with the APA's requirements for notice and comment where the use of traditional procedures are "impracticable, unnecessary, or contrary to the public interest."

regulations thus remain in effect during this period of review.[36] In the *Preliminary Results*, the Department used the average-to-transaction method for all U.S. sales to calculate the weighted-average dumping margin for Devi Fisheries. To comply with its regulations the Department must limit the application of the average-to-transaction methodology to only those Devi Fisheries' sales that have been found to be targeted.

### 2. **The Regulations Require That The Petitioners Allege Targeted Dumping, A Fundamental Requirement That Was Not Been Met In Many Cases**

The targeted dumping regulations state that Commerce "normally will examine only targeted dumping described in an allegation."[37] Allegations must be filed "no later than 30 days before the schedule date of the preliminary determination",[38] and must include "an explanation as to why the {A-T} or {T-T} method could not take into account any alleged price differences."[39] In this case, however, the petitioners have not made any allegations of targeted dumping.

Until the advent of the Cohen's d test, the Department had only initiated the targeted dumping investigation at the request of petitioners.[40] Indeed, in a number of cases the

---

[36] On April 22, 2014 Commerce published a Final Rule announcing its intention to not apply the withdrawn regulations to antidumping investigations.[36] Commerce maintains that the new Final Rule was promulgated in a manner consistent with requirements under the APA.[36] Even if Commerce were correct in this assertion, the scope of the Final Rule nevertheless does not cover the instant review because it only applies to cases "initiated on or after May 22, 2014."

[37] 19 CFR 351.414(f)(3)

[38] 19 CFR 351.301(d)(5)

[39] 19 CFR 351.414(f)(3).

[40] *Id.* at 1.

Department refused to initiate a targeted dumping investigation because the allegation was insufficient or unsubstantiated. [41] Such an approach is appropriate—the burden to identify and seek to redress targeted dumping should properly be upon the petitioners.

Since adopting the differential pricing analysis, the Department has self-initiated the targeted dumping investigation in many cases, without providing respondents any notice, to which they are entitled under the regulations. At a minimum, Commerce should be required to explain why it has chosen to investigate targeting dumping in many cases, despite the absence of allegations from the petitioners, when it "normally" requires specific allegations under its own regulations.

### 3. The Department's Current Methodolgy Does Not Establish That Significant Price Differences Exist

#### a. The Department 's Methodology Does Not Comply With Section 777A(d)(2)

The Department has stated that that the methodology for identifying targeted dumping in an administrative reviews parallels that applied in investigations and that Section 777A(d)(1) (B) should be instructive for purposes of examining whether to apply an alternative comparison methodology. [42] In providing notice that its default methodology in administrative reviews

---

[41] Unlike here, in the 2010-2011 review in Certain Frozen Warmwater Shrimp from India, the petitioners did allege targeted dumping for the two mandatory respondents. The petitioners alleged that significant price difference existed among purchasers and regions, but did not make any allegations concerning periods of time. The Department refused to initiate a targeting dumping allegation of Devi and Falcon in that previous review. See 78 Fed. Reg. 42492 (July 16, 2013).

[42] *See* Decision Memorandum at 5.

would be MA-MA, the Department indicated that it would "determine on a case-by-case basis whether it is appropriate to use an alternative comparison methodology by examining the same criteria that the Department examines in original investigations pursuant to sections 777A(d)(1)(A) and (B) of the Act." [43]

Commerce has not complied with the terms of Section 777A(d)(2), however. This statutory provision mandates Commerce to determine whether "there is a pattern of export prices (or constructed export prices) for comparable merchandise that **differ** significantly among purchasers, regions, or periods of time." The word "differ" is in the plural and hence relates to the export prices, not to the pattern. Therefore, the section requires Commerce to first determine whether the export prices for comparable merchandise differ significantly among purchasers, regions, or periods of time and then determine whether there is a pattern to export prices demonstrating such significant differences.

The differential pricing methodology calculates the weighted averages of prices for customers and regions for the entire period of investigation or review, and the periods by quarters and then tests for significant price differences. This is not in conformity with law because the price differences are not measured for export prices but they are measured for weighted averages of export prices. This leads to an enormous defect in the result because the averaged prices include individual export prices that show even insignificant price differences.

[43] *See Antidumping Proceedings: Calculation of the Weighted-Average Dumping Margin and Assessment Rate in Certain Antidumping Proceedings: Final Modification*, 77 Fed. Reg. 8101 (Feb. 14, 2012) ("Final Modification for Reviews.").

**b. The Term "Significant" Has Not Been Defined.**

Section 777A requires Commerce to find a "pattern of U.S. prices…that differ significantly among purchasers, regions or period of time."[44] It is critical to understand the meaning and implication of the word "significant" in order to implement the statute. The Merriam-Webster On-Line Dictionary defines the term "significant" as "of a noticeably or measureably large amount." Even in common parlance, the word "significant" implies something that is above normal. However, the Department considers all price differences within an averaged group to either pass Cohen d's test or not pass merely based on the weighted average. Thus all export prices, whether or not they display "significant" differences are treated as passing or not passing Cohen's d test. This flaw occurs because the differential pricing methodology fails entirely to give effect to the word "significant."

As described in greater detail below and in the attached statement from Dr. McFarland, the Cohen's d test upon which the Department relies is not designed to measure significant price differences, as required by the statute. Instead, it measures only the size of the difference between the means of two groups relative to the standard deviation. A difference that appears large in terms of the standard deviation may be very small in terms of actual prices.

The Department considers the Cohen's d statistic to indicate targeted dumping if it is greater than .8. The problem is that very small price difference may result in a Cohen's d statistic that is greater than .8. The Cohen's d test can produce a large positive result under

---

[44] 19 U.S.C. 1677f-1(d)(1)(B); see also 19 C.F.R. 351.414(f)(1)(i).

circumstances where the difference in prices is insignificant to the market but happens to exceed the standard deviation between the two values.[45]

From the legal standpoint, however, the critical issue is that the Department has failed to provide any guidance concerning the meaning of the term "differs significantly" as used in the targeted dumping provision. As the agency responsible for interpreting the statute, the Department has an obligation to define these terms. One of the justifications for proceeding on a case-by-case basis was explained in the SSA as follows:

> The Administration intends that in determining whether a pattern of significant price differences exist, **Commerce will 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.**[46]

In cases to date, the Department has not addressed the issue of what constitutes a significant difference in any particular industry. It has not examined the prices of individual transactions that passed the Cohen's d to determine whether the percentage differences between those prices and those that did not pass Cohen's d, are significant on the facts of the particular record before it. The Department's entire justification for conducting the targeted dumping analysis on a case-to-case basis, rather than on a uniform basis, is to examine the issues within the context of the industry being investigated or reviewed. Yet, there has been no attempt by the Department whatsoever to examine what level of differences in price will be significant for a particular industry.

---

[45] Exhibit A at 2

[46] SSA at 883

The Department is no stranger to the statutory term "significant" and has defined it in several other contexts. In Mid-Continent Nail Corp. v. United States,[47] the Court of International Trade upheld the Department's position that the term "differs significantly" means by five percent in the now discarded Nails test. Judge Restani noted that the five percent test has often been used to measure significance for antidumping purposes, and was not arbitrary and capricious.[48]

The Department recognized the need to define "differs significantly" for the Nails test. Yet, it has shirked its regulatory responsibility to define the term in the context of the current differential pricing methodology. The Department's refusal to interpret the statute that it is responsible for administering on the record of this case is both arbitrary and capricious.

Finally, the statute requires the Department to explain why "such differences cannot be taken into account" using the average-to-average methodology. See 19 U.S. C. 1677f-1(d)(1)(B)(ii). It is imperative to note that the term "such differences" relates to "export prices...that differ significantly" and not to the pattern, a factor that the current differential pricing methodology fails to take into account.

---

[47] 712 F. Supp. 2d 1370 (Ct. Int'l Trade 2010)

[48] *See, e.g.,* 19 U.S.C. §1677b(a)(1)(C)(1994) (using five percent test to determine home market viability);

19 U.S.C. §1677b(a)(1)(B)(ii)(II)(1994) (using five percent test to determine third-country market viability); 19 C.F.R. § 351.403(d)(1998) (using five percent test to determine whether to calculate normal value based on the sale of an affiliated party).

E. **<u>The Cohen's d Test is Seriously Flawed And Ill-Suited For Identifying Targeted Dumping.</u>**

1. **<u>The Cohen's *d* Test Fails to Identify Patterns of Significant Differences Among Prices Because It Was Meant To Measure Differencess In Well-Understood Units.</u>**

As an initial matter, Section 777A(d)(1)(B)(ii) requires the Department to explain "why such differences cannot be taken into account" using the average-to-average methodology. The differential pricing methodology as used by the Department ignores this entire legal mandate. First, the expression "such differences" is in the plural and thus relates to "export prices…that differ significantly" referred to in the preceding subparagraph. The Department should thus explain why the export prices that differ significantly cannot be taken into account by the A-A comparison method in investigations and the MA-T comparisons in reviews. The present differential pricing mechanism fails to do this.

The use of the Cohen's *d* statistic is arbitrary and capricious because it does not identify "pattern[s] of export prices"[49] for the purpose of "unmasking" targeted dumping. This is because the test is ill-suited to measuring differences in dollar-denominated transactions, and does not account for the commercial significance of any price differences it finds. Cohen's *d* is the wrong statistic to determine whether it is appropriate to use an alternative method.

The Cohen's d statistic was developed to measure "effect size" by Jacob Cohen, -- a professor of psychology-- and is the ratio of the difference between the means of two groups to their pooled

[49] *See* 19 USC 1677f-1(d)(B)(i); and 73 FR 74931 (noting Congress intended for A-T methodology to "unmask targeted dumping".

standard deviation, and is measured in terms of standard deviations. It was developed for use when actual units of measurement are difficult to understand or in comparing the results of studies that use different units of measure. Those circumstances are not present in a targeted dumping analysis because the phenomena being studied (prices) are measured in consistent well-understood units (dollars). In such circumstances, researchers will use actual units, rather than a measure where differences are expressed in terms of standard deviations, like Cohen's *d*.[50]

Cohen himself wrote that the use of his statistic:

> . . .proves salutary in those areas of the behavioral sciences where raw units are used which are quite arbitrary or lack meaning outside the investigation in which they are used, or both.[51]

He also wrote that if one is:

> comparing groups on a variable measured in units that are well understood by your readers (IQ points, **or dollars,** or number of children, or months of survival), mean differences are excellent measures of effect sizes. When this isn't the case, and it isn't the case more often than it is, the results can be translated into standardized mean differences (d values) or some measure of correlation or association.[52] (emphasis added ).

[50] Wilkinson, Leland and APA Task Force on Statistical Inference, Statistical methods in psychology journals" Guidelines and explanations," *American Psychologist* 54 (8) 1999: 594-604 at 599 ("If the units of measurement are meaningful on a practical level (*e.g.*, number of cigarettes smoked per day), then we usually prefer an unstandardized measure (regression coefficient or mean difference to a standardized measure (r or d).")

[51] Jacob Cohen, *Statistical Power Analysis for the Behavioral Sciences,"* second edition, Lawrence Erlbaum Associates, Hillsdale, NJ, 1988, p. 20.

[52] Jacob Cohen, "Things I have learned (so far)," *American Psychologist,* v. 45, no 12, December 1990, pp. 1304-1312.

Therefore Cohen himself never envisioned the test being used in circumstances such as targeted dumping analyses, when the comparison is between easily-understood units, because in these situations the test yields inaccurate results.

Ultimately, the test is useful for comparing the results of different studies that report results measured in different units or for analyzing results measured in units that are difficult to understand. Because that is not the case in the investigation of possible targeted dumping, it is inappropriate to use Cohen's *d* or other standardized measures of group mean differences.

**2. <u>The Use Of The .08 Benchmark Is Inappropriate</u>**

The Department calculates the Cohen's *d* statistic by first forming the test and base (or comparison) groups and then calculating the "*d*" statistic. Cohen suggested that a value of the statistic above .8 might in some cases be taken to be a large effect size.[53] Thus, if the statistic is above .8, the Department considers that the test group has passed the test, meaning it shows targeted dumping.

Though the Cohen's *d* statistic is used in a wide variety of contexts, and the use of the .8 benchmark may be appropriate in some of those contexts, its use is not appropriate in a test of targeted dumping. First, while the Department (in a prior investigation) cites a statistics book to the effect that this benchmark "has been widely adopted",[54] that same source calls the benchmark "somewhat arbitrary"

---

[53] Jacob Cohen, *Statistical Power Analysis for the Behavioral Sciences,* Second edition, Lawrence Erlbaum Associates, Hillsdale, NJ, 1988, p. 26.

[54] David Lane, et al., *Online Statistics Education: A Multimedia Course of Study,*(http://onlinestatbook.com/), p. 647. Cited in "Issues and Decision Memorandum for the Final Determination of the Antidumping Duty Investigation of Xanthan Gum from the People's Republic of China," U.S. Department of Commerce, International Trade Administration, May 28, 2013, p. 25.

and notes it has not been "universally adopted."[55] Second, the circumstances under which Cohen proposed .8 as a benchmark for classifying effect sizes as large under some circumstances were very different from those encountered in a study of targeted dumping. One article notes,

> Jacob Cohen defined effects as small (d=0.2), medium (d=.05), or large (d=.8). These rules of thumb were derived after surveying the behavioral sciences literature, which included studies in various disciplines involving diverse populations, interventions or content under study, and research designs. Cohen, in proposing these benchmarks in a 1988 test, explicitly noted that they are arbitrary and thus ought not to be viewed as absolute.[56]

Importantly, Cohen stated that these benchmarks were "recommended for use only when no better basis for estimating the ES [effect size] index is available."[57] Rules of thumb can hardly be considered substantial evidence.

Ultimately, the .8 benchmark has no relationship to the reasons for investigating potential targeted dumping, and therefore, its use in targeted dumping investigations is completely arbitrary and unjustifiable. Moreover, the use of that benchmark may make very small price differences seem to be "large effects." Thus, there is no justification for the use of the .8 benchmark in a study of targeted dumping.

---

[55] *Id.*

[56] Shayne B. Piasta and Laura M. Justice, "Cohen's *d* Statistic," *Encyclopedia of Research Design,* ed. Neil J. Salkind, Thousand Oaks Ca., SAGE Publications, 2010, p. 184.

[57] Jacob Cohen, *Statistical Power Analysis for the Behavioral Sciences,* second edition, Lawrence Erlbaum Associates, Hillsdale, N.J., 1988, p. 25.

The Cohen's *d* test is "considered complementary to the reporting of results from a test of statistical significance,"[58] meaning that the test must be interpreted in light of tests of statistical significance. In the absence of such interpretation, differences in averages among subgroups that are simply the result of random variation may be taken to indicate that a group is systematically charged lower or higher prices. The Department previously supported the absence of significance tests by citing an article that stated that a measure of effect size "has many advantages over the use of statistical significance tests alone."[59] The word "alone" is important in that citation. The article in fact says that effect size measures should be presented with tests of statistical significance. The article later states, "It is important to know the statistical significance of a result." It suggests that researchers "report the effect size, together with an estimate of its likely 'margin for error' or 'confidence interval.'"[60]

The Department also has argued that its "application of the Cohen's *d* test is based on the mean and variance calculated using the entire population of the respondent's sales in the U.S. market, and, therefore, these values contain no sampling error. Accordingly, statistical significance is not a relevant

---

[58] Shayne B. Piasta and Laura M. Justice, "Cohen's *d* Statistic," op. cit., p. 180.

[59] David Coe, "It's the Effect Size Stupid," Paper presented at the Annual Conference of the British Educational Research Association, September 2002, p. 1, available at http://www.leeds.ac.uk/educoUdocuments/00002182. Cited in "Issues and Decision Memorandum for the Final Determination of the Antidumping Duty Investigation of Xanthan Gum from the People's Republic of China," U.S. Department of Commerce, International Trade Administration, May 28, 2013, p. 24.

[60] David Coe, "It's the Effect Size Stupid," Paper presented at the Annual Conference of the British Educations Research Association, September 2002, p. 5.

consideration in this context."[61] Besides being an incorrect test, because the Cohen's $d$ statistic is measured in terms of standard deviation, it can be difficult to interpret the statistic. The variance and thus the standard deviation of a group are affected by how the group is defined. Because of the way it is constructed, Cohen's $d$ statistic is not useful for finding meaningful price differences.

The Department often bases groups on strictly defined homogeneous products, meaning that many of the subgroups the Department uses in the Cohen's $d$ test are homogenous and have very low standard deviations. Because of this, differences that appear large relative to a low standard deviation may be very small in terms of actual price differences. The smallest price difference that would be found to be a "large effect" under the Cohen's $d$ test will depend on the ratio of the base (comparison) group mean to the pooled standard deviation (pooled across the test and comparison group). The higher that ratio, the lower the smallest price difference found to be a large effect. This point can be shown by a few simple equations. Suppose that the comparison group mean is M, and the percentage difference between the comparison and test group means is x, so the test group mean is (1+x) M. The pooled standard deviation is PSD. Then the Cohen's $d$ statistic is

$$(M-(1+x)M)/PSD.$$

A group is considered to have passed the Cohen's $d$ test if the statistic's absolute value is greater than .8, that is if

$$|\ (M-(1+x)M)/PSD\ | > .8.$$

That equation can be simplified to

---

[61] U.S. Department of Commerce International Trade Administration, "Issues and Decision Memorandum for the Final Results of Antidumping Duty Administrative Review: Polyethylene Terephthalate Film From India; 2011-2012 Administrative Review," February 21, 2014, p. 9.

$|(-x*M)/PSD| > .8$.

As the mean and pooled standard deviation are always positive, that equation can be expressed as

$|x| > .8/(M/PSD)$.

That is, Cohen's *d* test will be passed anytime the percentage difference between the test and comparison group mean prices exceeds .8 divided by the ratio of the mean to the pooled standard deviation, the MSD ratio.

Many comparison groups used in a antidumping case may be very homogeneous, so the MSD ratio is often very large. Thus, the Cohen's *d* test often is calculated under conditions such that a group could pass the test even though there is very little difference between its mean and the base group mean.

Cohen's *d* test is the wrong test for investigating significant price differences indicating possible targeted dumping. The test uses standardized units, which are not preferred because the measured units, dollars, are well understood. Moreover, in this instance, standardized units produce highly misleading results by magnifying small differences in prices.

### 3. <u>The Cohen's *d* Test Is Likely To Find False Positives</u>

The Cohen's *d* test used in antidumping determinations may indicate targeted dumping for a subgroup, even though there is no attempt to target particularly low or high prices to the region, quarter, or purchaser associated with that subgroup. Thus, the test has a high probability of finding targeted dumping and thus is subject to finding "false positives."

**a. The Cohen's *d* Test Uses Improper Criteria For Finding Patterns Of Pricing Differences**

The Department includes transactions as exhibiting apparent targeted dumping even when the overall test results for a purchaser, region, or period fail to demonstrate a pattern of high or low prices. However, test groups should only be included in the group that may demonstrate targeted dumping if the purchaser, region or period test groups being considered to pass the test all show high prices or all show low prices.[62]

The selection criteria used by the Department in an antidumping determinations are also subject to another form of inconsistency. Many transactions that supposedly pass the Cohen's *d* test do so for inconsistent reasons. Transactions are classified into subgroups that are defined by a control number, a region, a period, and a purchaser (The control number is determined by the specific Product involved). Subgroups are aggregated into test groups over two of the last three characteristics and then tested for targeted based on the third. Thus, each subgroup of transactions is tested three times.[63] In some instances, a subgroup may be considered as demonstrating possible targeted dumping because the subgroup was found to have high prices in one test and low prices in another. Such conflicting test results should not be taken to show targeted dumping. This type of inconsistency arises because some subgroups are considered to

[62] This procedure still overstates how many transactions may show targeted dumping because the procedure uses test groups even where many other test groups for the purchaser, region, or period being considered do not pass the test.

[63] There is a test based on region, against a comparison aggregation by quarter and purchases; a test based on quarter, against a comparison aggregation by region and purchaser; and a test based on purchaser, against a comparison aggregation based on region and quarter.

have exhibited possible targeted dumping, even though results are inconsistent across different tests involving the same subgroup.

In looking for targeted dumping, the Department is supposedly looking for patterns of pricing differences. The Cohen's *d* test does not do that not only because the Cohen's *d* is the wrong statistic to use but also because the Department uses improper criteria for determining whether a group exhibits apparent targeted dumping.

**b. Improper Comparison Groups And Circularity**

Furthermore, the Department's differential pricing methodology is circular. The methodology begins by isolating a set of data from the total dataset, for use as a test group. The remaining data then becomes the base group. The test group is tested against base group. Afterwards, another set of data is isolated from total data as the test group and the remaining total data *including* the first test group now becomes the base data. This methodology creates a circular calculation, and is thus fundamentally flawed. After a test group is found to have unusually high or low prices, it is still included in comparison groups to determine if other test groups have high or low prices. Suppose for example, that the test indicates unusually high prices in the West region. Those prices are then included in a comparison group and used to test whether prices in the Northeast, Midwest, and South are different from those in other regions. Because the comparison group for each of those regions includes the unusually high prices in the West, prices in each of the other regions may be found to be unusually low, thus "passing" the test, even though these prices in fact do not differ from those charged in other regions.

This circularity, however, was not always an issue. The model margin calculation program on the Department's website in October 2012 stated as follows:

> US-13-D-ii. Begin the DP test. After each round (region, time, then customer), sales that pass the DP test will be removed from the DPSALES database. Those that were either not tested, or were tested but failed, will remain in (or be put back into) DPSALES and pass on to the next round of testing.

Further, in the *Nails* test used by the Department prior to the differential pricing analysis, there was no such circular calculation. Specifically in the Nails test, the petitioners alleged TD in respect to specified customers and periods, which were called 'Alleged Targets' ('AT'). The rest of the periods and customers were 'Non Targets' ('NT'). The ATs constituted the numerator and the NTs constituted the denominator in the calculations.

Thus antidumping determinations made by the Department are based in part on contradictory procedures. Some prices are found to be anomalous, in the sense that they are higher or lower than those for other regions, periods, or purchasers. Then those anomalous prices are included in multiple comparison groups to determine if other prices are unusually high or low. This circular approach leads to unjustifiable results. Consider the following example: during the period of review there were imports of Product X only in the third and fourth quarters. However, the volume in the fourth quarter was more than twice that in the third quarter and the fourth quarter average price was lower, but only by 1.7% or just over 5 cents. When the test group was formed for the third quarter, the comparison group was only comprised of the fourth quarter. While the prices in the two quarters were very close, the pooled standard deviation was very low, .0050. Thus, the Cohen's d statistic was -10.22, and the test group passed. Then the test group was formed for the fourth quarter, and the comparison group was the third quarter. Thus, essentially the same test was run twice. The Cohen's d statistic was 10.22, and the fourth quarter passed the test. Even though there is little difference between the prices in the different quarters, both quarters were classified as showing targeted dumping. Both quarters cannot be "targeted" and the result is double counting of sales that passed Cohen's d. Thus, the Cohen's *d* test is inherently flawed in its propensity to yield false positives.

## III. CONCLUSION

For the reasons we have stated herein, we believe the Cohen's *d* test is flawed and we urge the Department to discontinue its application of the test in investigating price differences that might indicate targeted dumping.

Respectfully submitted,

Ronald M. Wisla

Lizbeth R. Levinson for
Kutak Rock LLP

Counsel to the Seafood Exporters' Association of India

# EXHIBIT A

Economists
INCORPORATED
Washington DC • San Francisco

*In affiliation with*
The Allen Consulting Group, Australia

2121 K Street, NW
Suite 1100
Washington, DC 20037

Henry B. McFarland
Direct Dial: (202) 833-5262
mcfarland.h@ei.com

Main (202) 223-4700 Fax (202) 296-7138

June 22, 2014

# Statement of Henry B. McFarland

## 1. Introduction

I am an economic consultant with the firm of Economists Incorporated, where I have worked since 1989. Prior to joining that firm, I worked as an economist with the Antitrust Division of the U.S. Department of Justice for over eight years and with the U.S. International Trade Commission for over three years. My work as an economist has often involved the use of statistical analyses to investigate various economic questions. I have a Ph.D. in Economics from Northwestern University. My curriculum vitae is attached to this statement as Exhibit 1.

Counsel has asked me to consider a number of economic issues pertaining to the Department of Commerce's (the Department's) differential pricing analysis. In writing this report, I am relying on my background and knowledge as an economist. I am also relying on various publicly available sources, and on the Department's decisional memoranda and the data used by the Department in its preliminary determination in the recent administrative review concerning shrimp from India. (referred to henceforth as "preliminary determination")[1]

## 2. Summary of Conclusions

I have reached several major conclusions, which are described in this report. My main conclusions are as follows:

- Cohen's d test was designed for specific purposes and is not a proper test for investigating price differences that might indicate targeted dumping.

1 U.S. Department of Commerce, International Trade Administration, "Decision Memorandum for the Preliminary Results of the 2012-2013 Administrative Review of the Antidumping Duty Order on Certain Frozen Warmwater Shrimp from India," March 18, 2014 ("preliminary determination").

- The Cohen's d statistic is measured in terms of standard deviations: It was developed for use when actual units of measurement are difficult to understand or in comparing the results of studies that use different units of measure. Those circumstances are not present here.

- A difference that appears large in terms of the standard deviation may be very small in terms of actual prices. Thus, Cohen's d test does not indicate the commercial significance of the price differences it finds.

- The Department considers Cohen's d statistic to indicate targeted dumping if its absolute value is greater than .8. The .8 benchmark does appear in the statistics literature, but only in contexts that are very different from those found in an investigation of targeted dumping. Use in those contexts does not justify its use here.

- The use of the .8 benchmark in targeted dumping investigation is arbitrary and unjustifiable. That benchmark has no relationship to the reasons for investigating potential targeted dumping. Moreover, the use of that benchmark may make very small price differences seem to be "large effects." Thus, there is no justification for the use of that benchmark in a study of targeted dumping.

* The Cohen's d test uses improper criteria for finding patterns of pricing differences. These criteria are subject to two forms of inconsistency in identifying transactions that may exhibit targeted dumping.

  - The Department includes transactions as passing the test even when the overall results do not show a pattern of high or low prices for the purchaser, region, or quarter being considered. For example, test groups involving a specific purchaser may be classified as showing targeted dumping even though some tests involving the purchaser show high prices, some show low prices, and many show no difference.

  - Many transactions that supposedly pass the Cohen's d test do so for inconsistent reasons; one test shows them to be high-priced and another shows them to be low-priced. Transactions are classified into subgroups that are defined by a control number, which is determined by the product, a region, a quarter, and a purchaser. Subgroups are aggregated over two of the last three characteristics to form test groups. Those test groups are then tested for targeted dumping based on the third characteristic. Thus, each subgroup of transactions is grouped into three different test groups and tested three times. A number of subgroups are considered to have passed the Cohen's d test because they had prices that were high based on one test but low based on another test. If one test finds high prices and another finds low prices, that is no indication of a pattern of high or low prices.

- The Cohen's d test may indicate targeted dumping for a subgroup, even though there is no attempt to target particularly low or high prices to that group.
    - As the Cohen's d test is currently performed, it does not test any specific hypothesis of targeted dumping; rather it simply tests for differences between the means of various groups of prices.
    - A transaction may be included in the group exhibiting apparent targeted dumping on any one of three grounds: region, quarter, or purchaser. Thus, it has three chances of having prices that are classified as either too high or too low. *Many of the transactions that pass the Cohen's d test may pass only one of the three times that they are subjected to the test.*
    - A test group that is found to have unusually high or low prices is still used in comparison groups to determine if other test groups have high or low prices. Prices that are found to be anomalous are included in multiple comparison groups to determine if other prices are unusually high or low.
    - The Department's current methodology does not include a test of statistical significance even though the statistics literature states that the Cohen's d test is a complement to such tests. Thus, differences in averages that are simply the result of random variation may be taken to indicate that a group is systematically charged lower or higher prices.
- Even if one were to accept the proposed justification for the use of the average-to-transaction method including some zeroing (that the high prices on some import sales may obscure the margins emanating from low prices on other sales) that still would not justify all uses of zeroing.
    - Using the Department's decision rules, a large share of transactions may have dumping margins calculated using zeroing even though they do not pass Cohen's d test. If more than 66% of sales pass Cohen's d test, the Department

uses the average-to-transaction method with zeroing for 100% of sales. Based on that rule, zeroing will be used for transactions that did not pass the Cohen's d test, as long as they are less than 34% of sales.

- The Department uses zeroing for transactions that have passed the Cohen's d test because they are in groups that have average prices *below* those of the comparison group. As these groups have been found to have low prices, there is no reason to believe that high prices charged to them obscure dumping that favors other groups. Thus, they should not be included in the group for which zeroing is used.

- In addition, with the mixed method, if the aggregate dumping margin for the transactions that are subject to the average-to-average calculation is negative (higher US average prices), then that aggregate is increased to zero. There is no justification for that step, as those are the very transactions that have specifically been found not to involve targeted dumping.

## 3. Cohen's d test is inappropriate for investigating targeted dumping.

### a. The Cohen's d test was developed for other purposes and cannot be used for this purpose.

The approach to targeted dumping used by the Department is based on Cohen's d statistic. That statistic, which was developed to measure "effect size" by Jacob Cohen, a professor of psychology, is the ratio of the difference between the means of two groups to their pooled standard deviation. The Department first forms test and base (or comparison) groups and then calculates the d statistic. Dr. Cohen suggested that a value of the statistic above .8 might in some cases be taken to be a large effect size.[2] Thus, if the absolute value of the statistic is above .8, the Department considers that the test group has passed the test, and shows targeted dumping.

Cohen's d test, however, was developed for different purposes than those for which the Department is using it, and its use here is improper. As noted, Cohen's d statistic measures differences between group means in terms of standard deviations. Thus, it is useful for comparing the results of different studies that report results measured in different units or for analyzing results measured in units that are difficult to understand.[3] Because that is not the case

---

[2] Jacob Cohen, *Statistical Power Analysis for the Behavioral Sciences,* second edition, Lawrence Erlbaum Associates, Hillsdale, N.J., 1988, p. 26.

[3] As an example of units that are hard to understand, Cohen gives a measure of people's attitude towards the United Nations where the units are points with a higher number of points signifying a more favorable attitude. It

in the investigation of possible targeted dumping, it is inappropriate to use Cohen's d or other standardized measures of group mean differences.

In the analysis of targeted dumping, the phenomena being studied, prices, are measured in consistent well-understood units, dollars. In those circumstances, researchers generally prefer to use actual units, rather than a measure where differences are expressed in terms of standard deviations, like Cohen's d.[4] Cohen himself wrote that the use of his statistic

> ". . .proves salutary in those areas of the behavioral sciences where raw units are used which are quite arbitrary or lack meaning outside the investigation in which they are used, or both."[5]

He also wrote that if one is

> "comparing groups on a variable measured in units that are well understood by your readers (IQ points, **or dollars**, or number of children, or months of survival), mean differences are excellent measures of effect sizes. When this isn't the case, and it isn't the case more often than it is, the results can be translated into standardized mean differences (d values) or some measure of correlation or association."[6] [emphasis added]

Because the Cohen's d statistic is measured in terms of standard deviation, it may be difficult to interpret the statistic. The variance and thus the standard deviation of a group are affected by how the group is defined. In the preliminary determination, the Department used a largely arbitrary method of grouping the data for the test. For example, it used five quarters for the periods, but it could as well have used months or seasons. It divided the continental United States into four Census regions: Northeast, South, Midwest, and West. But the Census Bureau subdivides those 4 regions into 9 smaller regions, and the Department could as well have used those. Thus, the ratio of the group mean to the standard deviation is determined in part by arbitrary decisions concerning how the Cohen's d test is to be run. That ratio has no economic significance. Purchasers do not choose their sources by comparing differences in price to the

---

would be unclear how much difference in attitude is measured by a point. Jacob Cohen, *Statistical Power Analysis for the Behavioral Sciences,* second edition, Lawrence Erlbaum Associates, Hillsdale, N.J., 1988, p. 20.

4 Wilkinson, Leland and APA Task Force on Statistical Inference, "Statistical methods in psychology journals: Guidelines and explanations," *American Psychologist* 54 (8) 1999: 594–604 at 599. ("If the units of measurement are meaningful on a practical level (e.g., number of cigarettes smoked per day), then we usually prefer an unstandardized measure (regression coefficient or mean difference) to a standardized measure (r or d).")

5 Jacob Cohen, *Statistical Power Analysis for the Behavioral Sciences, "* second edition, Lawrence Erlbaum Associates, Hillsdale, N.J., 1988, p. 20.

6 Jacob Cohen, "Things I have learned (so far)," *American Psychologist*, v. 45, no. 12, December 1990, pp. 1304-1312.

standard deviation of prices. Nor is the standard deviation of prices an input into the calculation of the dumping margin.

Nor is the use of the .8 benchmark appropriate for a test of targeted dumping. In the preliminary determination, the Department considers Cohen's d statistic to indicate targeted dumping if its absolute value is greater than .8. The Department previously cited a statistics book to the effect that this benchmark "has been widely adopted."[7] That same source calls the benchmark "somewhat arbitrary" and notes it has "not been universally accepted."[8] Cohen's d test is used in a wide variety of contexts, and the use of the .8 benchmark may be appropriate in some of those contexts. It is not appropriate here.

Cohen proposed .8 as a benchmark for classifying effect sizes as large under some circumstances, but those were very different from those encountered in a study of targeted dumping. One article notes,

> "Jacob Cohen defined effects as small (d=0.2), medium (d=.5), or large (d=.8). These rules of thumb were derived after surveying the behavioral sciences literature, which included studies in various disciplines involving diverse populations, interventions or content under study, and research designs. Cohen, in proposing these benchmarks in a 1988 test, explicitly noted that they are arbitrary and thus ought not to be viewed as absolute."[9]

Cohen stated that these benchmarks were "recommended for use only when no better basis for estimating the ES [effect size] index is available."[10]

The use of the .8 benchmark in this matter is completely arbitrary and unjustifiable. That benchmark has no relationship to the reasons for investigating potential targeted dumping. Moreover, as discussed in more detail below, the use of that benchmark may make very small price differences seem to be "large effects." Thus, there is no justification for the use of that benchmark in a study of targeted dumping.

---

7 David Lane, et al., *Online Statistics Education: A Multimedia Course of Study*, (http://onlinestatbook.com/), p. 647. Cited in "Issues and Decision Memorandum for the Final Determination of the Antidumping Duty Investigation of Xanthan Gum from the People's Republic of China," U.S. Department of Commerce, International Trade Administration, May 28, 2013, p. 25.

8 David Lane, et al., op. cit.

9 Shayne B. Piasta and Laura M. Justice, "Cohen's d Statistic," *Encyclopedia of Research Design*, ed. Neil J. Salkind, Thousand Oaks Ca., SAGE Publications, 2010, p. 184.

10 Jacob Cohen, *Statistical Power Analysis for the Behavioral Sciences,* second edition, Lawrence Erlbaum Associates, Hillsdale, N.J., 1988, p. 25.

Because of the way it is constructed, Cohen's d statistic is not useful for finding meaningful price differences. The Department has based groups on strictly-defined homogeneous products. Thus, many of the test and comparison groups the Department uses in the Cohen's d test are homogenous and have very low standard deviations. A difference that appears large relative to a low standard deviation may be very small in terms of actual price differences.

The smallest price difference that would be found to be a "large effect" under the Cohen's d test will depend on the ratio of the base (comparison) group mean to the pooled standard deviation (pooled across the test and comparison group). The higher that ratio, the lower the smallest price difference found to be a large effect. This point can be shown by a few simple equations. Suppose that the comparison group mean is M, and the percentage difference between the comparison and test group means is x, so the test group mean is (1+x)M. The pooled standard deviation is PSD. Then the Cohen's d statistic is

$$(M-(1+x)M)/PSD.$$

A group is considered to have passed the Cohen's d test if the statistic's absolute value is greater than .8, that is if

$$|(M-(1+x)M)/PSD| > .8.$$

That equation can be simplified to

$$|(-x*M)/PSD| > .8.$$

As the mean and pooled standard deviation are always positive, that equation can be expressed as

$$|x| > .8/(M/PSD).$$

That is, Cohen's d test will be passed anytime the percentage difference between the test and comparison group mean prices exceeds .8 divided by the ratio of the mean to the pooled standard deviation, the MSD ratio.

Many comparison groups used in an antidumping investigation are very homogeneous, so the MSD ratio is often very large. Thus, the Cohen's d test often is calculated under conditions such that a group could pass the test even though there is very little difference between its mean and the base group mean. In the recent administrative review involving shrimp from India, for example, one group had a Cohen's d statistic of -2.42 and thus passed the Cohen's d test by a wide margin. The difference between the mean of that test group and its comparison group was only .49%, less than half of one percent, of the comparison group mean. Another group passed the Cohen's d test with a test statistic of .94. The difference between the mean of that test group and its comparison group, however, was only .03%.

Cohen's d test is the wrong test for investigating significant price differences indicating possible targeted dumping. The test uses standardized units, which are not preferred because the measured units, dollars, are well understood. Moreover, in this instance, standardized units produce highly misleading results by magnifying small differences in prices.

### b. The Cohen's d test uses improper criteria for finding patterns of pricing differences.

In looking for targeted dumping, the Department is supposedly looking for patterns of pricing differences. The Cohen's d test conducted in the preliminary determination does not do that not only because the Cohen's d is the wrong statistic to use but also because the Department uses improper criteria for determining whether a group exhibits apparent targeted dumping.

The Department includes transactions as exhibiting apparent targeted dumping even when the overall test results for a particular purchaser, region, or period do not show a pattern of high or low prices. Changing only this part of the procedures used would have a dramatic effect on the results. Test groups should only be included in the group that might show evidence of targeted dumping if the test groups for the purchaser, region, or period being considered that pass the test all show high prices or all show low prices.[11]

The selection criteria used by the Department are also subject to another form of inconsistency. Many transactions that supposedly pass the Cohen's d test do so for inconsistent reasons. Transactions are classified into subgroups that are defined by a control number, a region, a period, and a purchaser. (The control number is determined by the specific product involved.) Subgroups are aggregated into test groups over two of the last three characteristics and then tested for targeted dumping based on the third. Thus, each subgroup of transactions is tested three times.[12] In some cases, a subgroup may be considered to exhibit possible targeted dumping because it was found to have high prices in one test and low prices in another test. Such conflicting test results should not be taken to show targeted dumping.

[11] This procedure still overstates the number of transactions that might show targeted dumping because it includes test groups even if many other test groups for the purchaser, region, or period being considered fail the test and for other reasons to be discussed later.

[12] That is, there is a test based on region, against a comparison aggregation by quarter and purchaser; a test based on quarter, against a comparison aggregation by region and purchaser; and a test based on purchaser, against a comparison aggregation based on region and quarter.

This form of inconsistency is different from that discussed above. The first type of inconsistency arises because different test groups are considered to have exhibited possible targeted dumping, even though results are inconsistent across different test groups pertaining to the same purchaser, period, or region. This second type of inconsistency arises because some subgroups are considered to have exhibited possible targeted dumping, even though results are inconsistent across different tests involving the same subgroup.

If one test finds high prices and another finds low prices, that is no indication of a pattern of high or low prices for that subgroup. Thus, such subgroups should be excluded from the share of transactions considered to have passed Cohen's d test.

## 4. The Cohen's d test is likely to find false positives.

The Cohen's d test may indicate targeted dumping for a subgroup, even though there is no attempt to target particularly low or high prices to the region, quarter, or purchaser associated with that subgroup. There are several reasons why the test has a high probability of finding targeted dumping and thus is subject to finding "false positives."

As noted above, a transaction may be included in the group subject to targeted dumping on any one of three grounds: region, period, or purchaser. Thus, it has three chances of having prices that are classified as either too high or too low. Most of the transactions that pass the Cohen's d test pass only one of the three times that they are subjected to the test.

### a. Improper comparison groups

Also after a test group is found to have unusually high or low prices, it is still included in comparison groups to determine if other test groups have high or low prices. Suppose for example, that the test indicates unusually high prices in the West region. Those prices are then included in a comparison group and used to test whether prices in the Northeast, Midwest, and South are different from those in other regions. Because the comparison group for each of those regions includes the unusually high prices in the West, prices in each of the other regions may be found to be unusually low, thus "passing" the test, even though these prices in fact don't differ from those charged in other regions.

Thus, the Department's methodology is based in part on a self-contradictory procedure. Some prices are found to be anomalous, in the sense that they are higher or lower than those for other regions, periods, or purchasers. Then those anomalous prices are included in multiple comparison groups to determine if other prices are unusually high or low.

Table 1 below shows a simple example that illustrates how this aspect of the procedure used in the preliminary determination may affect the results. In the example, the West is assumed to have higher average prices. The other three regions have the same prices. To simplify, every test and base (comparison) group is assumed to have the same standard deviation. The test is run first on the prices in the West, and the test is passed. When the test is run on the Northeast, it is also passed, because the absolute value of the Cohen's d statistic is greater than .8. The test is passed even though the Northeast has the same prices as the other two regions because the comparison group includes the low prices in the West. The same is true when the test is performed for the South and Midwest. Thus, even though only one region actually has higher prices, all four regions are included in the value that passes the Cohen's d test. If the West were excluded from the comparison group, none of the other regions would pass the test.

**Table 1: An Example of the Effects of Using Anomalous Groups in the Base Group**

| | Test Group Average Price | Base Group Average Price | Pooled Standard Deviation | Cohen's d |
|---|---|---|---|---|
| West | $7.00 | $5.00 | 0.25 | 8.00 |
| Northeast | $5.00 | $5.67 | 0.25 | -2.67 |
| South | $5.00 | $5.67 | 0.25 | -2.67 |
| Midwest | $5.00 | $5.67 | 0.25 | -2.67 |

This problem could be ameliorated if the Department would not consider both high and low-priced groups as passing the test. As will be discussed later, applying the average-to-transaction method to the low-priced groups is not consistent with supposed justifications for that method.. Thus, the Department could consider only high-priced test groups, those with average prices above those of the comparison group, as passing the test. In that case, having groups that passed the test because they had anomalously high prices in the comparison group might make other groups look anomalously low, but it would not cause them to pass the test.

As will be discussed later, eliminating the test groups with low prices from the groups that are classified as showing targeted dumping will greatly reduce the share of the value of imports that appear to pass the test.

### b. Lack of tests for statistical significance

Moreover, the reporting of a Cohen's d statistic is "considered complementary to the reporting of results from a test of statistical significance."[13] That is, the d statistic must be interpreted in light of tests of statistical significance. Nonetheless, the Department does not test

[13] Shayne B. Piasta and Laura M. Justice, "Cohen's d Statistic," op. cit., p. 180.

for statistical significance. Thus, differences in averages between test and comparison groups that are simply the result of random variation may be taken to indicate that a group is systematically charged lower or higher prices.

The Department previously supported the absence of significance tests by citing an article that stated that a measure of effect size "has many advantages over the use of statistical significance tests alone."[14] The word alone is important in that citation. The article in fact says that effect size measures should be presented with tests of statistical significance. The article later states, "It is important to know the statistical significance of a result." It suggests that researchers "report the effect size, together with an estimate of its likely 'margin for error' or 'confidence interval.'"[15]

The Department also has argued that its "application of the Cohen's *d* test is based on the mean and variance calculated using the entire population of the respondent's sales in the U.S. market, and, therefore, these values contain no sampling error. Accordingly, statistical significance is not a relevant consideration in this context."[16] In fact, no Cohen's d test is based on the entire population of the respondent's sales. Before tests are run, all transactions are broken down into small groups defined by specific products.

### c. Failure to specify and test a specific hypothesis

Finally, as the Cohen's d test is currently performed by the Department, it does not test any specific hypothesis of targeted dumping involving any particular group of purchasers or region or period; rather it simply tests for differences between the means of various groups. The Department has simply applied the Cohen's d test seriatim to a variety of groups, to test whether prices to some groups were different (either lower or higher) than prices to others. Because the Department is not testing a specific hypothesis, the group definitions are arbitrary. There is no reason to use Census regions as opposed to states or quarters as opposed to months or weeks. That procedure raises the possibility that random fluctuations in price will be mistaken for targeted dumping.

---

[14] David Coe, "It's the Effect Size Stupid," Paper presented at the Annual Conference of the British Educational Research Association, September 2002, p. 1, available at http://www.leeds.ac.uk/educol/documents/00002182. Cited in "Issues and Decision Memorandum for the Final Determination of the Antidumping Duty Investigation of Xanthan Gum from the People's Republic of China," U.S. Department of Commerce, International Trade Administration, May 28, 2013, p. 24.

[15] David Coe, "It's the Effect Size Stupid," Paper presented at the Annual Conference of the British Educational Research Association, September 2002, p. 5.

[16] U.S. Department of Commerce International Trade Administration, "Issues and Decision Memorandum for the Final Results of Antidumping Duty Administrative Review: Polyethylene Terephthalate Film From India; 2011-2012 Administrative Review," February 21, 2014, p. 9.

## 5. Conclusion

The Cohen's d test would be the wrong test to use in the context of targeted dumping. Moreover, the Department currently uses the wrong criteria for determining what test groups should be considered to exhibit possible targeted dumping. Test groups that do not show a pattern of either high or low prices should not be classified as exhibiting apparent targeted dumping. Finally, even if the Department does choose to use an average-to-transaction methodology with some zeroing, it should only use zeroing in instances that are consistent with its justification for the practice.