

United States Court of Appeals
FOR THE DISTRICT OF COLUMBIA CIRCUIT

Argued October 21, 2010

Decided December 28, 2010

No. 10-1019

UNION PACIFIC RAILROAD COMPANY,
PETITIONER

v.

SURFACE TRANSPORTATION BOARD AND UNITED STATES OF
AMERICA,
RESPONDENTS

US MAGNESIUM, LLC,
INTERVENOR

On Petition for Review of an Order
of the Surface Transportation Board

Jonathan Marcus argued the cause for petitioner. With him on the briefs were *J. Michael Hemmer*, *Louise A. Rinn*, and *Michael L. Rosenthal*.

James A. Read, Attorney, Surface Transportation Board, argued the cause for respondent. With him on the brief were *Robert B. Nicholson* and *John P. Fonte*, Attorneys, U.S. Department of Justice, *Craig M. Keats*, Deputy General Counsel, Surface Transportation Board, and *J. Frederick Miller*,

Jr., Attorney. *Ellen D. Hanson*, General Counsel, entered an appearance.

Thomas W. Wilcox and *David K. Monroe* were on the brief for intervenor US Magnesium, L.L.C. in support of respondent.

Before: GRIFFITH and KAVANAUGH, *Circuit Judges*, and EDWARDS, *Senior Circuit Judge*.

Opinion for the Court filed by *Senior Circuit Judge EDWARDS*.

EDWARDS, *Senior Circuit Judge*: US Magnesium, L.L.C. (“USM”) operates a magnesium production facility in Rowley, Utah, and relies on the Union Pacific Railroad Company (“UP”) to ship its chlorine co-product from Rowley to two receiving facilities in Arizona. In May 2009, USM filed a complaint before the Surface Transportation Board (“STB” or the “Board”), arguing that UP’s rates for the two chlorine shipments, also called “movements,” were unreasonably high. USM opted to bring its challenge under the Three Benchmark framework set forth in *Simplified Standards for Rail Rate Cases*, STB Ex Parte No. 646 (Sub-No. 1), 2007 WL 2493509 (STB served Sept. 5, 2007), *aff’d sub nom. CSX Transp., Inc. v. STB*, 568 F.3d 236 (D.C. Cir. 2009), *vacated in part on other grounds on reh’g*, 584 F.3d 1076 (D.C. Cir. 2009). Under this framework, both the shipper and the rail carrier submit groups of comparison rates (or “offers”) for the movements in question, from which the Board makes an “either/or” choice, with no modifications allowed once the final submissions have been made. In January 2010, after oral arguments and the submission of opening evidence, reply evidence, and rebuttal evidence from both parties, the Board selected USM’s comparison rate groups. Using the USM groups, the Board found UP’s rates for the two chlorine movements to be unreasonable. *US Magnesium, L.L.C. v. Union Pac. R.R.*, STB Docket No. 42114, 2010 WL 319727

(STB served Jan. 28, 2010) (“*STB Decision*”), reprinted in Joint Appendix (“J.A.”) 823-47. UP now petitions for review of the Board’s decision, arguing that the Board’s selection of USM’s comparison rate groups was arbitrary and capricious and seeking vacatur of the judgment and remand for a re-application of the Three Benchmark framework using UP’s proffered comparison rate group.

We find no grounds to reverse the Board’s decision, particularly given the deference owed to the Board in rate-making disputes. Under the Three Benchmark framework, the Board was obliged to choose between two positions, neither of which was ideal. USM’s comparison rate groups were problematic, because they were primarily comprised of anhydrous ammonia shipments. UP’s comparison rate group was flawed, because it included an unduly heavy sample of rebilled traffic. However, based on the quantitative data and the residual differences between the comparison rate groups, the Board determined that USM’s submission was more representative of the ideal comparison group – namely, single-line chlorine traffic. Because the Board’s decision “articulated a rational connection between the facts found and the decision made,” *N. Am. Freight Car Ass’n v. STB*, 529 F.3d 1166, 1170-71 (D.C. Cir. 2008) (quoting *PPL Mont., LLC v. STB*, 437 F.3d 1240, 1245 (D.C. Cir. 2006) (internal quotation marks omitted)), we deny the petition for review.

I. BACKGROUND

A. *The Three Benchmark Framework for Rate Reasonableness*

By statute, rail carriers are authorized to engage in a certain amount of demand-based differential pricing in order to earn “adequate revenues,” 49 U.S.C. § 10701(d)(2), whereby “captive” shippers who are more dependent on rail service due to the lack of transportation alternatives may be charged a higher

markup as compared to “competitive” shippers who are free to pursue lower-cost transportation options. *See BNSF Ry. v. STB*, 526 F.3d 770, 776 (D.C. Cir. 2008) (noting that rail carriers would lose revenue if they imposed a strict *pro rata* share of the joint and common costs to each shipper regardless of the shipper’s degree of captivity). However, in order to protect captive shippers from shouldering an unconscionable share of carriers’ general operations costs, Congress has provided that, in situations when “a rail carrier has market dominance over the transportation to which a particular rate applies, the rate established by such carrier for such transportation must be reasonable.” 49 U.S.C. § 10701(d)(1). In reviewing rate challenges, the Board must therefore consider whether carriers are maximizing revenue from competitive shippers, *id.* § 10701(d)(2)(B), and whether “one commodity is paying an unreasonable share of the carrier’s overall revenues,” *id.* § 10701(d)(2)(C).

Mindful of the high litigation costs associated with full stand-alone cost (“SAC”) presentations in cases involving rate challenges, Congress instructed the Board to “establish a simplified and expedited method for determining the reasonableness of challenged rail rates in those cases in which a full stand-alone cost presentation is too costly, given the value of the case.” *Id.* § 10701(d)(3); *see also Simplified Standards* at 5 (commenting that recent full-SAC presentations had incurred nearly \$5 million in litigation expenses and that even simplified-SAC presentations could run to a cost of \$1 million). Pursuant to this congressional mandate, the Board adopted the Three Benchmark framework, a relatively straightforward, “final offer” methodology that measures the reasonableness of a rate by comparing the revenue-over-variable cost (“R/VC”) ratio of the challenged rate to the R/VC ratios of a comparison group of similar movements. The challenged rate is held presumptively unreasonable if it has a higher R/VC ratio than the calculated upper boundary of the comparison group. *See Simplified*

Standards at 21-22 & n.30 (explaining that the upper boundary is the 90 percent confidence interval from a one-sided hypothesis test constructed around the mean of the comparison group).

The Three Benchmark framework allows both the rail carrier and the challenging shipper to submit a comparison rate group, with the only requirement being that all movements included must be “captive movements.” *Id.* at 17 (explaining that a captive movement is defined as a movement with a R/VC ratio greater than 180 percent, meaning that the revenue produced from shipping the movement was enough to cover 100 percent of the variable costs and still left a surplus in the amount of 80 percent of the variable costs to be applied towards the rail carrier’s joint and common costs). The movements are drawn from an unmasked sample of a carrier’s Waybill data – a comprehensive database that includes detailed information on all of the movements shipped by a rail carrier, including origin and destination points, type of commodity shipped, tonnage shipped, and revenue. There are multiple rounds of revisions, during which each party is given an opportunity to adjust its proposed comparison rate group. After the parties submit their final offers, the Board then selects the comparison rate group that is “most similar in the aggregate to the [challenged] movement[.]” *Id.* at 18.

In making its selection, the Board considers “a variety of factors, such as length of movement, commodity type, traffic densities of the likely routes involved, and demand elasticity.” *Id.* at 17. While “movements with different cost characteristics may be included in the comparison group,” the Board “will favor a comparison group that consists of movements of like commodities.” *Id.* To encourage both parties to offer competitive and reasonable comparison groups, the Board’s selection is an “either/or” choice between the parties’ final offers, with no modifications allowed. *Id.* at 18.

After a comparison rate group is selected, the R/VC ratio of each movement in that chosen group is then multiplied by the carrier's "revenue need adjustment factor," a figure that serves to "reflect demand-based differential pricing principles." *STB Decision* at 12-13, J.A. 834-35. If the challenged rate's R/VC ratio is greater than the upper boundary of the adjusted comparison group, the rate will be presumed unreasonable. Since the revenue need adjustment factor is derived from static figures published annually by the Board, the Three Benchmark framework's reasonableness determination generally turns on the Board's selection of a comparison group. The shipper bringing the rate challenge will naturally have an economic incentive to pick movements with low R/VC ratios while the rail carrier defending the reasonableness of the rate will have an equal incentive to pick comparison movements with high R/VC ratios. Although the Board recognizes that the Three Benchmark framework has some inherent limitations, neither party challenges the validity of the framework itself in this case.

B. *Factual History and Proceedings Before the Board*

USM operates the nation's only primary magnesium production facility in Rowley, Utah, and generates chlorine as a co-product of the magnesium production process. Due to chlorine's status as a toxic inhalation hazard ("TIH"), USM relies on UP to transport the chlorine as liquefied compressed gas by rail service to two receiving facilities in Eloy and Sahuarita, Arizona. In March 2009, USM began shipping its chlorine under UP's common carrier tariff rates, having been unable to secure an extension of the previous contract rates. UP charged USM \$13,396 per carload for shipments from Rowley to Eloy, and \$10,410 per carload for shipments from Rowley to Sahuarita, a marked increase from the previous years' contractually negotiated rates. Based on variable costs of \$2,549 and \$2,485, respectively, the resulting R/VC ratios were 526 percent for Eloy and 419 percent for Sahuarita. *STB*

Decision at 3, J.A. 825. USM filed a rate challenge with the Board in May 2009, opting to use the Three Benchmark framework established in *Simplified Standards*. Since UP acknowledged its market dominance over the chlorine movements at issue, the Board immediately moved to the comparison group selection process.

The Eloy and Sahuarita movements at issue are chlorine shipments traveling in privately owned tanker cars with capacities of less than 22,000 gallons for 1,290 and 1,250 load miles, respectively, on single-line car service. USM's opening evidence proposed two separate comparison groups, one for the Eloy movement and one for the Sahuarita movement, including movements of all TIH commodities but limiting the groups to movements on single-line service within a range of plus or minus 200 miles of the actual rail distances. *STB Decision* at 6, J.A. 828. UP submitted a single 24-movement comparison group for both the Eloy and Sahuarita movements, including both single-line and rebilled traffic and increasing the mileage band to 400 miles but limiting the movements to those of chlorine-only traffic and those traveling in cars that are smaller than 22,000 gallons in capacity. *Id.* at 6-7, J.A. 828-29.

Both USM and UP challenged the other's proposed groups. Among other things, UP argued that USM did not select "like commodities" because its groups were comprised, in large part, of anhydrous ammonia movements and wrongly included movements traveling in cars larger than 22,000 gallons. USM countered that UP's group had too few movements, wrongly included rebilled traffic, and was over-inclusive due to the larger 400-mile range. The Board agreed that the comparison rate groups submitted by both parties were flawed.

All else being equal, local single-line chlorine movements would be the preferable comparison group for the issue movements. However, neither party presented that comparison group. Rather, both parties chose relatively

extreme comparison groups in their initial tenders. And while the deficiencies in each of the comparison groups were evident at opening, and either party could have strengthened its selection by adopting movements from the other's group(s), both parties stood rigidly behind their initial, deficient comparison group selections.

Id. at 9, J.A. 831.

The Board's analysis focused on the two most "pivotal" differences: USM's inclusion of non-chlorine movements and UP's inclusion of rebilled traffic. *Id.* at 7-9, J.A. 829-31. It acknowledged UP's argument that anhydrous ammonia had significantly different demand characteristics and lower transportation risks as compared to chlorine, thus distorting the comparability of USM's proposed groups. *Id.* at 7 & nn.9-10. However, it found UP's inclusion of rebilled traffic similarly problematic, noting that the R/VC ratios for the rebilled movements in the UP comparison group were much higher than those of the single-line movements. *Id.* at 8, J.A. 830. Faced with two imperfect offers, the Board attempted to compare the distortions. The Board found that single-line anhydrous ammonia had an average R/VC ratio 19 percent less than single-line chlorine and that rebilled chlorine had an average R/VC ratio 31 percent greater than single-line chlorine. *Id.* at 9-10, J.A. 831-32. When these distortions were weighted based on their proportional representation in the parties' comparison groups, USM's Eloy and Sahuarita comparison groups had a distortion effect of 16 percent and 17 percent, respectively, while UP's comparison group had a distortion effect of 18 percent. *Id.* The Board conceded that this difference in weighted impacts "may not be 'statistically significant,'" but concluded that the analysis "confirms what USM has advocated, which is [that] significant bias [was] introduced when [UP] included rebilled traffic in its comparison group." *Id.* at 10, J.A. 832.

Having decided that single-line, chlorine traffic would be the “preferable comparison group,” *id.* at 9, J.A. 831, the Board noted that the R/VC ratios of USM’s comparison groups were “nearly identical to the markups of the single-line, chlorine traffic in UP’s group.” *Id.* at 11, J.A. 833 (finding adjusted R/VC ratio for USM’s groups to be 304 percent and 298 percent as compared to 301 percent for UP’s single-line chlorine movements while UP’s rebilled movements had an R/VC ratio of 475 percent). The Board referenced the residual differences between the groups, finding that USM’s groups were strengthened by a “larger sample size and narrower mileage bands, which are more reasonable with respect to the length of haul than UP’s mileage bands because they are closer to the two issue movements’ actual miles.” *Id.* at 10, J.A. 832.

The Board then offered this telling explanation:

In sum, based on our quantitative analysis and the residual differences between the groups, we find that USM’s comparison groups provide the best evidence of a reasonable level of contribution to joint and common costs for the issue movements. . . .

[B]ased on this record, USM’s comparison groups provide a better gauge of the *proper* comparison group. The traffic at issue constitutes single-line, chlorine movements. Yet UP submitted a comparison group containing 42% single-line traffic and 58% rebilled (i.e., interchanged) traffic. Single-line movements are different from rebilled movements. Single-line movements travel over a single carrier, and in this case an average distance of just over 1,000 miles. Rebilled traffic, on the other hand, is carried by multiple carriers with multiple interchanges and handoffs, and the total distance traveled by the commodity typically will be much longer. As such, the costs of transporting the commodity are different. Further, the markup over variable cost UP will charge for its portion of the total trip will not

only be a function of the shipper's elasticity of demand for transportation, but also a function of UP's relative bargaining power compared with other carriers involved in the movement.

When we scrutinize UP's comparison group, the difference in the average adjusted R/VC ratios between single-line and rebilled movements is significant: 301% for single-line chlorine traffic in UP's comparison group compared to 475% for rebilled traffic in UP's comparison group. UP was unable to justify allowing rebilled traffic to skew its comparison group. Moreover, as USM stressed at our oral argument, its comparison groups actually provide R/VC levels nearly identical to the markups of the single-line, chlorine traffic in UP's group. The adjusted R/VC ratios submitted by USM for the Eloy and Sahuarita movements were 304% and 298%, compared to 301% for single-line, chlorine movements in UP's group. . . .

To address differences in both demand elasticities and transportation risks, nothing prevented UP from providing a comparison group containing just single-line, chlorine movements, and excluding rebilled movements. We appreciate why it chose not to: it would have won the battle (the selection of the comparison group) and lost the war (the result would be the same as in this case). If UP had provided a comparison group of single-line, chlorine movements, the issue of how our costing model treats movements of chlorine would be largely irrelevant, because this is a *comparison* approach, and the issue movements and comparison movements would be treated similarly by our costing model. See *Simplified Standards*, at 17, 84.

Id. at 11, J.A. 833 (emphases in original). The Board ultimately chose USM's comparison rate groups, stating that "USM's submission (notwithstanding the relative lack of chlorine traffic)

provides more reasonable comparison groups than one so sharply skewed by rebilled traffic.” *Id.* at 12, J.A. 834.

Using USM’s groups for the remainder of the Three Benchmark analysis, the Board found the maximum reasonable R/VC ratio to be 356 percent for the Eloy movement and 346 percent for the Sahuarita movement. *Id.* at 13-19, J.A. 835-41. As both challenged rates had R/VC ratios that exceeded these upper boundaries, the Board ordered UP to establish and maintain new rates within the allowable range and to reimburse USM for amounts previously collected above the prescribed levels. *Id.* at 19, J.A. 841.

Commissioner Nottingham dissented from the decision, focusing on the lack of chlorine movements in USM’s comparison groups and questioning the majority’s reliance on rebilled traffic’s distortive effect without a clear “understanding of the reasons for the apparent difference between the R/VC ratios of single-line vs. rebilled chlorine movements,” particularly when USM had the burden as the challenging party of showing unreasonableness. *Id.* at 24-25, J.A. 846-47. The dissent criticized the majority’s use of a narrow and self-created quantitative analysis to “break the tie.” *Id.* at 24, J.A. 846. UP timely petitioned for review of the Board’s decision, and USM intervened.

II. ANALYSIS

A. *Standard of Review*

This court reviews decisions of the Board with deference. A Board decision will not be reversed unless it is “unsupported by substantial evidence,” 5 U.S.C. § 706(2)(E), “arbitrary, capricious, an abuse of discretion, or otherwise not in accordance with law,” *AEP Tex. N. Co. v. STB*, 609 F.3d 432, 438 (D.C. Cir. 2010); 5 U.S.C. § 706(2)(A). Deference is particularly high in rate disputes, where “the Board acts at the zenith of its powers.” *Id.* (quoting *Burlington N. R.R. v. STB*,

114 F.3d 206, 210 (D.C. Cir. 1997) (internal quotation marks omitted). “Where the Board’s findings rest on such relevant evidence as a reasonable mind might accept as adequate to support a conclusion, and where the Board has articulated a rational connection between the facts found and the decision made, [this court] will not disturb its judgment.” *N. Am. Freight Car Ass’n*, 529 F.3d at 1170-71 (quoting *PPL Mont.*, 437 F.3d at 1245 (internal quotation marks omitted)).

B. UP’s Challenge Fails

UP’s challenge to the Board’s decision is premised on its belief that its proposed comparison rate group, even with its heavy inclusion of rebilled traffic, was not skewed. UP maintains that the Board’s decision is arbitrary and capricious, because it considers rebilled traffic on par with non-chlorine traffic without explaining *why* rebilled traffic would have a distortive effect. UP’s argument is unpersuasive, for it is grounded on a serious misunderstanding of the Three Benchmark framework.

The Three Benchmark, final-offer process does not facilitate a “search for the truth.” The Board’s final judgment in any case involving this final-offer process necessarily will be constrained by what the parties present as their final offers. The Board merely selects between two final offers, neither of which may be ideal. There is no room for compromise, and no procedure to adjust two flawed final offers.

It is true that USM, as the party bringing the rate challenge, had the ultimate burden of persuasion to show that UP’s rates were unreasonable. See *BNSF Ry. v. STB*, 453 F.3d 473, 485 (D.C. Cir. 2006). However, in the comparison group selection stage of the Three Benchmark analysis, USM and UP each had an obligation to defend its own proposal against the opposing side’s attacks. See *id.* (citing “*Complex*” *Consol. Edison Co. v. FERC*, 165 F.3d 992, 1008 (D.C. Cir. 1999)) (explaining that the

shipper bringing the challenge who holds the ultimate burden of persuasion does not necessarily have the burden of production “[s]o long as the record supports [the Board’s ultimate] conclusion”). There is nothing in this record to indicate that UP’s proposed comparison group, and the resulting maximum rate level, would have been judged objectively reasonable under a full-SAC rate challenge. Indeed, even UP does not make this claim. Thus, the question before the Board was which of the proposed comparison rate groups was least undesirable. This is the nature of the beast in the “final offer” process upon which the Three Benchmark framework is founded. UP’s protests ignore this reality.

UP challenged the comparability of anhydrous ammonia to chlorine traffic, arguing that USM’s heavy inclusion of anhydrous ammonia skewed the demand characteristics of its comparison rate group. USM tried to defend its selections by referencing the Board’s stated preference for comparison groups with a mix of TIH commodities in *E.I. DuPont de Nemours and Co. v. CSX Transportation, Inc.*, STB Docket No. 42100, 2008 WL 2588609 (STB served June 30, 2008), the only previous application of the Three Benchmark framework. *STB Decision* at 7-8, J.A. 829-30. However, the Board pointed out that *DuPont* is easily distinguishable. *Id.* (explaining that *DuPont*’s carrier had confessed to overt chlorine pricing manipulation). USM, in turn, challenged the comparability of rebilled traffic to single-line traffic, arguing that UP’s heavy inclusion of rebilled traffic skewed the R/VC ratios of its own comparison group. Like USM, UP then had the opportunity to explain or refute the alleged distortion. Like USM, UP failed to make a convincing defense.

UP acknowledged that rebilled traffic has lower variable costs as compared to single-line traffic. While a difference in variable costs, by itself, does not affect the comparability of two movements, the presumption that comparability is not affected

depends on the existence of similar R/VC ratios. *See Simplified Standards* at 17 (noting that movements with different costs are acceptable because “there is no reason, *a priori*, to presume that the R/VC ratios (or their share of joint and common costs) should be different”). *Here, there was undisputed evidence that the rebilled traffic in UP’s sample had much higher R/VC ratios as compared to single-line traffic.* The difference in variable costs was, therefore, not matched by similarly lowered prices, but rather by a greater profit margin, suggesting differences in demand elasticity under a differential pricing system.

Building upon USM’s allegation that rebilled movements had much higher R/VC ratios, the Board compared the relative distortive impact of anhydrous ammonia and rebilled movements, and found that rebilled chlorine movements had an average R/VC ratio 31 percent higher than single-line chlorine movements, but that anhydrous ammonia movements had an average R/VC ratio only 19 percent lower than chlorine movements. *STB Decision* at 9-10, J.A. 831-32. Though it acknowledged that the weighted impact between the two distortions might “not be ‘statistically significant,’” the analysis “confirm[ed] what USM ha[d] advocated, which is [that] significant bias [was] introduced when [UP] included rebilled traffic in its comparison group.” *Id.* at 10, J.A. 832.

At oral argument before the Board, counsel for UP was specifically asked to explain the difference caused by the inclusion of rebilled traffic, *see Oral Arg. Tr.* 28, Nov. 23, 2009, J.A. 743 (“How do you respond to [USM’s] argument regarding the re-bill issue in his exhibit?”). UP offered no explanation, either in its rebuttal evidence or in its response to the Board’s questioning. UP’s only response was to maintain that even if the use of rebilled traffic was distortive, USM’s inclusion of non-chlorine traffic was even more damaging:

[If] it’s a choice between re-billed movements with some cost characteristics that might be slightly different and a

comparison group that's comprised 99 percent or 96 percent of non-chlorine traffic, where should you be more concerned that you're not reflecting the demand characteristics? I submit it's the comparison group that has hardly any chlorine, not the fact that you've got some re-billed movements in there.

Id. at 35, J.A. 750. Since "UP was unable to justify allowing rebilled traffic to skew its comparison group," *STB Decision* at 11, J.A. 833, the Board reasonably concluded based on the evidence presented that rebilled traffic, like anhydrous ammonia, had a comparably improper distortive effect.

Since both UP and USM's groups were flawed, the Board proceeded to compare both groups to the ideal comparison group – namely, single-line chlorine movements. The Board explained why USM came out ahead in the final-offer process:

In the end, it is our responsibility to select the better comparison group between those submitted by the parties in order to judge the reasonableness of the challenged rates. . . . Here, we agree with UP and our dissenting colleague that there are important differences between movements of chlorine and anhydrous ammonia, as illustrated by UP's evidence of the differences between the two groups. While that difference, as measured by the R/VC ratios in the broader Waybill sample data, is significant (19%), a more pronounced difference is present between single-line and rebilled, chlorine traffic (31%). These simplified proceedings do not offer us a proper platform to explore why either discrepancy exists. But all of the quantitative evidence points toward accepting USM's groups over that proffered by UP. Given that: (1) UP's own comparison group shows average R/VC levels for single-line, chlorine movements that are practically identical to those from USM's groups; (2) the supplemental quantitative analysis we undertook of the relevant data confirms that the

distortion from rebilled movements appears more significant; and (3) the residual differences between the parties' groups – such as the number of observations and the mileage bands – also favor USM's groups, it is our judgment that USM's submission (notwithstanding the relative lack of chlorine traffic) provides more reasonable comparison groups than one so sharply skewed by rebilled traffic.

Id. at 12, J.A. 834.

It is important to say again that the Board's judgment need not reflect mathematical certainty in order to survive judicial scrutiny. Indeed, there is good reason to believe that judgments rendered pursuant to the Three Benchmark framework more often than not will be the antithesis of mathematical certainty. The final-offer process does not offer or promise certainty or precision in ratemaking. It would be absurd for anyone to think otherwise. Therefore, because, as we have already noted, "the Board acts at the zenith of its powers when it sets rail rates, we recognize that the Board is entitled to particular deference. As long as the Board has rationally set forth the grounds on which it acted, this court may not substitute its judgment for that of the agency." *AEP Tex. N. Co.*, 609 F.3d at 438 (citations, ellipsis, and internal quotation marks omitted). The Board's decision easily meets this standard.

UP disregards the reality of the situation here and criticizes the Board's emphasis on single-line chlorine traffic as the ideal comparison group. UP argues that the Board should have considered whether rebilled movements may actually be more comparable. This argument is illogical: why would the Board have found rebilled traffic more comparable than single-line traffic when the two movements in question were single-line movements? UP also criticizes the Board's emphasis on its "supplemental quantitative analysis," arguing that a test resulting in statistically insignificant weighted differences cannot be dispositive. However, the Board did not say that the

quantitative analysis was dispositive. The Board's comparison of R/VC distortion was merely one portion of its overall analysis. Finally, UP criticizes the Board's reliance on the residual differences, but it cannot dispute the fact that USM's mileage bands were more representative of the two movements at issue. Where, as here, the Board's selection was based on a rational application of evidence presented, this court has no power to "substitute its [own] judgment for that of the agency." *Id.* (quoting *BNSF Ry.*, 453 F.3d at 480).

The Three Benchmark framework was created to give shippers a less costly vehicle through which to bring their rate disputes. The process was designed to give parties an incentive to advance proposals that would minimize their differences, in search of reasonable accommodations. Unfortunately, both USM and UP chose to take extreme positions in their comparison rate group offers. The final-offer process thus failed to work as intended in this case.

The Board has explained how the process should work:

Any final tender that is skewed too far in one direction might well result in the selection of a more reasonable final tender presented by the opposing party. By having two rounds of simultaneous tenders and a technical conference, both sides will participate in the winnowing process. . . . This approach will work as intended only if the parties know that the agency will not attempt to find a compromise position somewhere in the middle. To create the proper incentives for the litigants not to take extreme positions, we commit to selecting the more reasonable of the two groups as tendered.

Simplified Standards at 18. Here, the anticipated "winnowing process" did not occur. Both sides stood firmly behind their initial, clearly skewed offers, apparently hoping that the other side's proposal would be rejected as even more extreme.

Indeed, “[t]he parties seem[ed] to be intent on testing the outer boundaries of what might qualify as an acceptable comparison group, rather than adhering to the spirit of the Three-Benchmark process and seeking middle ground in an effort at reaching a reasonable and expedited outcome.” *STB Decision* at 22 (Commissioner Nottingham, dissenting), J.A. 844. UP lost this high-stakes gamble and now protests that the result is unjust. But UP cannot blame the Board for this result. The Board’s decision, being rationally based on the facts presented, was not in error.

III. CONCLUSION

For the foregoing reasons, we deny the petition for review.