United States Court of Appeals FOR THE DISTRICT OF COLUMBIA CIRCUIT Argued January 15, 2002 Decided March 1, 2002 No. 01-1066 Association of Oil Pipe Lines, Petitioner v. Federal Energy Regulatory Commission and United States of America, Respondents Canadian Association of Petroleum Producers, et al., Intervenors On Petition for Review of an Order of the Federal Energy Regulatory Commission C. Frederick Beckner III argued the cause for petitioner. With him on the briefs were Lawrence A. Miller and Michele F. Joy. Dennis Lane, Solicitor, Federal Energy Regulatory Commission, argued the cause for respondents. With him on the brief was Cynthia A. Marlette, Acting General Counsel. John J. Powers III and Robert J. Wiggers, Attorneys, U.S. Department of Justice, entered appearances. James H. Holt and Melvin Goldstein were on the brief for intervenors. Before: Sentelle and Rogers, Circuit Judges, and Williams, Senior Circuit Judge. Opinion for the Court filed by Senior Circuit Judge Williams. Williams, Senior Circuit Judge: In December 2000 the Federal Energy Regulatory Commission established a formu- la for changes in the ensuing years' price caps for interstate oil pipelines. See Order Concluding Initial Five-Year Re- view of the Oil Pipeline Pricing Index, 93 FERC p 61,266 (2000) ("Order" or "2000 Order"). To drive the annual change in the caps, it chose the Producer Price Index for Finished Goods minus one percent ("PPI-1"). Petitioner Association of Oil Pipe Lines challenges this as arbitrary and capricious, saying that the FERC Staff report justifying continuing adherence to the PPI-1 index used statistical methods that deviated from FERC's previous methodology without apparent justification, and that it also failed to ac- count for special factors potentially altering the pattern of future changes. We find FERC's responses to the Associa- tion's criticisms inadequate--except as to the special fac- tors--and therefore remand for further proceedings. * * * In prior orders FERC adopted a price cap regime for oil pipelines. See Revisions to Oil Pipeline Regulations Pursu- ant to the Energy Policy Act of 1992, Order No. 561, F.E.R.C. Stats. & Regs. (CCH) p 30,985 (1993) ("Order No. 561"); Revisions to Oil Pipeline Regulations Pursuant to Energy Policy Act of 1992, Order No. 561-A, F.E.R.C. Stats. & Regs. (CCH) p 31,100 (1994) ("Order No. 561-A"); see also Association of Oil Pipe Lines v. FERC, 83 F.3d 1424, 1429- 30 (D.C. Cir. 1996) ("AOPL I"). After fixing as a baseline the pipeline rates that Congress deemed "just and reasonable" in the Energy Policy Act of 1992, Pub. L. No. 102-486, 106 Stat. 2776, 3010 (1992) ("EPAct"), reprinted in 42 U.S.C. s 7172 note, FERC determined to use an indexing scheme to make annual adjustments. The index initially picked was PPI-1. FERC said that it was making this choice because, when compared to various alternatives, PPI-1 seemed to most closely track historical changes in actual pipeline costs. Order No. 561 at 30,951/2. But FERC's choice of PPI-1 was not "for all time." Order No. 561-A at 31,092-93. To ensure continuing fit between the index and actual changes in industry costs, FERC as- sured commentators that it would reexamine the index every five years. Order No. 561 at 30,941/2. In 2000 FERC embarked on the first such reexamination. In its Notice of Inquiry it cited a Staff study purporting to show that "the changes in the PPI-1 Index have closely approximated the changes in the reported cost data for the oil pipeline industry during the five-year period covered by [the] review." Notice of Inquiry, Five-Year Review of Oil Pipeline Pricing Index, 65 Fed. Reg. 47,358, at 47,361 (2000) (report- ing Staff study results). FERC invited comments. The Association responded, claiming that the Staff study deviated from past methodology and was otherwise flawed. Com- ments of the Association of Oil Pipe Lines, Five-Year Review of Oil Pipeline Pricing Index, Docket No. RM00-11-000 (Sept. 1, 2000) ("AOPL Comments"). It argued that the FERC Staff had improperly measured cost changes, had erroneously failed to remove statistical outliers, and had inexplicably altered its method for calculating capital costs. And it said that the Staff had failed to account for factors that would likely cause future cost changes to diverge from the historical trend. In fact, it said, PPI was a more appropriate index than PPI-1. The Commission rejected the Association's arguments and issued the order now under review. See 28 U.S.C. s 2344. 1. Measurement of Cost Changes. The Association's first contention is that FERC used an

improper methodology in pursuing its stated intention to measure "actual cost changes experienced by the oil pipeline industry." Order at 61,849/1; Order No. 561-A at 31,092/2; see also Order No. 561 at 30,952/2. Many methods are available. One possibility is to calculate the percentage cost change (per barrel-mile) for each individual firm and combine them in a simple average. Another is to combine the firm barrel-mile costs in an aver- age weighted by volume, so that minor firms do not skew the result. Another is to take the median of the distribution. We will refer to these methods respectively as the unweight- ed average, the fixed-weight average, and the median. As we shall see, there are other candidates as well. Orders Nos. 561/561-A substantially cited and relied on a study that reported the results of all three of the methods described above, as well as a composite figure that combined the three. See Test. of Alfred E. Kahn, Revisions to Oil Pipeline Regulations Pursuant to the Energy Policy Act of 1992, Docket No. RM93-11-000 (Aug. 12, 1993), at 11 tbl.1 ("1993 Kahn Study"). The 1993-94 orders do not unambigu- ously show which figure held a dominant position in FERC's reasoning. The change in the composite for each of the three periods considered was fairly close to PPI-1. Of the three types of averages making up the composite, the unweighted average was closest to the composite (and thus to the PPI-1 figure ultimately selected).1 _____ 1 The 1993 Kahn Study, upon which Orders Nos. 561/561-A substantially relied, reported the following:

	1982-87	1987-92	1982-92
Operating expenses and net plant			
Weighted average	0.82%	2.49%	1.24%
Unweighted average	0.11%	1.27%	1.54%
Median	-0.26%	0.45%	0.85%
Composite	0.22%	1.40%	1.21%
Producer Price Index	1.06%	3.17%	2.11%
Difference from composite	0.84%	1.77%	0.90%

1993 Kahan Study, at 11 tbl.1. In any event, we need not determine FERC's precise method in 1993 because the current order uses none of these previous methods. Instead of calculating cost changes by individual firm and then averaging them by any of the methods used before, the 2000 FERC Staff report used what we may call a "floating-weight" average. For each year in the period 1994-99 it took total costs for the entire industry, and divided it by the total number of barrel-miles shipped, yielding an annual average industry cost per barrel-mile. This produced an annual change, and the study found these annual changes to be a bit lower on average than the annual change in PPI-1. Notice of Inquiry, 65 Fed. Reg. 47,359-60 & tbl.1. We call this method a "floating-weight" average because it effectively weights each pipeline's per-barrel costs by that pipeline's volume. In contrast, a fixed-weight average weights each firm's cost change by the firm's market share (in either the previous year or the current year). As was shown by the pipelines' expert witness, Professor Alfred E. Kahn (interestingly, the expert relied on by the shippers in the 1993-94 round), a floating-weight average can yield odd results. One curiosity, for example, is that such an average will include the costs of new entrants, even though, not having been in the market, they will have experienced no "change" in cost at all. More generally, changes in market share among partici- pants can give an arguably distorted impression of cost changes. Professor Kahn offered the following example: Suppose in Year 1, pipeline A's costs are $2 per barrel-mile, and its volume is 5 barrel-miles. Pipeline B's costs are $0.50 per barrel-mile, and its volume is 2 barrel-miles. In Year 2, B's volume remains the same, but A's volume decreases to 4. In addition, from Year 1 to Year 2, both pipelines experience an increase in cost by 12.5% (i.e., their respective costs per barrel-mile increase from $2 to $2.25 for pipeline A and from $0.50 to $0.5625 for pipeline B). Under a fixed-weight aver- age, the average cost change is plainly 12.5%. Under a floating-weight average, however, the calculated change in cost is only 7.6%.2 See Kahn Decl., Five Year Review of Oil Pipeline Pricing Index, Docket No. RM00-11-000 (Aug. 31, 2000), at 7 ("2000 Kahn Study"). One can, of course, produce a more extreme hypothetical by adjusting the numbers, creat- ing a scenario where all pipelines experience a uniform in- crease in costs but the floating-weight average shows a de- cline. FERC does not deny these peculiarities. Instead, it makes several collateral arguments in support of its approach, all of which are unpersuasive. First it responds that the Associa- tion's claims of underestimation are simply the consequence of competition and the

move by customers from higher-cost providers to lower-cost providers. Order at 61,850 (arguing that the Association's results are "simply the natural working of the market forces at play, and does not show any distortion resulting from Staff's methodology"). Continuing, FERC argues that the Association's fixed-weight approach "would raise the price ceiling and thereby enable more high-cost pipelines to become or remain profitable." Id. The problem with FERC's "competition" theory is that even if it were sound as a general matter (and FERC makes no effort to vindicate it), it presumes that all pipelines in the industry are close substitutes for each other. But by all indications in the record, they are not. Cf. Farmers Union Central Exchange, Inc. v. FERC, 734 F.2d 1486, 1508 n.50 (D.C. Cir. 1984) (agreeing with the Justice Department that competition in the oil pipeline industry "must be evaluated in terms of discrete regional markets"). Indeed, if there were close competition between the pipelines, the reason for rate regulation--each pipeline's market power--would be missing. See id. at 1508 ("It is of course elementary that market failure and the control of monopoly power are central ratio- _____ 2 This figure is calculated as follows: Aggregate industry cost in Year 1 is [($2.00)*5 + ($0.50)*2] / 7 = $1.57 per barrel-mile. Aggregate industry cost in Year 2 is [($2.25)*4 + ($0.5625)*2] / 6 = $1.69 per barrel-mile. The percentage change is ($1.69-$1.57) / $1.57 = 7.6%. nales for the imposition of rate regulation.") (citing Stephen Breyer, Regulation and Its Reform 15-16 (1982)). Moreover, Professor Kahn offered data--uncontested by FERC--suggesting that the oil pipeline industry is divided between relatively high-cost crude oil pipelines ($0.013 per barrel-mile in 1999) and relatively low-cost "product" pipe- lines (carrying refined products) ($0.0038 per barrel-mile). See Kahn Reply Decl., Five-Year Review of Oil Pipeline Pricing Index, Docket No. RM00-11-000 (Oct. 2, 2000), at 9. Indeed, both Professor Kahn and the expert for the shippers, Professor Frederic M. Scherer, refer to crude and product pipelines as separate markets. Id. (discussing how demand for crude pipelines has declined, whereas it has risen for product pipelines); Scherer Decl., Five-Year Review of Oil Pipeline Pricing Index, Docket No. RM00-11-000 (Aug. 28, 2000), at 3-4 (noting that the two types of pipelines "face rather different demand conditions", and that there has been "conversion of some crude pipelines into product pipelines"). If this is true, it seems entirely likely that the movement from higher-cost crude to lower-cost product pipelines has relative- ly little to do with competition between them. In a similar vein, FERC suggests that its approach (with its resulting choice of a lower index) "emphasizes the ... efficiency-promoting (i.e., cost controlling) property ... of using an indexing system." Order at 61,849/1. There are two possible aspects to this assertion. First, FERC may be claiming that its approach creates cost-controlling incentives. But the general theory behind price caps is that because they are largely disconnected from individual firm costs, a firm is not dissuaded from cost-cutting efforts by the prospect of seeing its ceiling lowered. Order No. 561 at 30,948-49 n.37; see also Nat'l Rural Telecom Ass'n v. FCC, 988 F.2d 174, 178 (D.C. Cir. 1993). It is that disconnect, rather than the method of calculating price caps, that creates the good incen- tives. As Professor Kahn observed (partly facetiously), one could achieve the purpose by "relat[ing] the change in permis- sible prices over time to a random table of numbers." 1993 Kahn Study, at 4. The second possible meaning of FERC's efficiency- promoting argument might be called the "survival" incentive. The prospect of imminent bankruptcy surely concentrates the mind. But if this is the justification, it amounts to no more than the principle that "lower is better"--an argument that seems to have no end and little connection to any stated purpose. FERC further responds in its brief that even if the Associ- ation's criticisms of its methodology are valid, its members can always resort to a safety valve built into FERC's rate adjustment scheme. Specifically, under Orders Nos. 561/561-A, a pipeline can file for "cost-of-service" rates based on itsindividualized costs if it "can demonstrate that there is a substantial divergence between the actual costs experienced by the pipeline and the indexed ceiling rate." Order at 61,850 n.35; see also 18 C.F.R. s 342.4; AOPL I, 83 F.3d at 1430- 31. But by definition, a "safety valve" should only address aberrant cases, however broadly

this class may be defined. See Order No. 561-A at 31,106-07 (expanding the circum- stances under which cost-of-service ratemaking is permitted); Order No. 561 at 30,956-57 (same). A safety valve cannot rescue FERC's indexing methodology from systemic errors, for then the exception would swallow the rule. See Time Warner Entertainment Co. v. FCC, 56 F.3d 151, 173 (D.C. Cir. 1995) (refusing to allow FCC to invoke a cost-of-service option, "a limited 'safety-valve' exception," to justify its deci- sion). Furthermore, as FERC acknowledges, cost-of-service rate filings are "more cumbersome" and "more costly and time-consuming." Respondent's Brief at 17. And a regime based in large part on their use would be inconsistent with Congress's mandate under the EPAct for FERC to establish "a simplified and generally applicable ratemaking methodolo- gy." EPAct, at s 1801(a). Finally, FERC's counsel at oral argument sought to deflect Professor Kahn's criticisms by seeming to characterize him as some kind of economic Svengali. Counsel said that he was "perfectly confident [that] no matter what method [FERC] had used, if Doctor Kahn wanted to come up with examples to make it look silly, he could have." Oral Argument Tr. at 27. However flattering this may be to Professor Kahn, it is hardly a defense for FERC. An expert's acknowledged skill is no basis for refusing to confront his analysis. The Commission's Order mentions a possible defense to its use of the "floating weight" methodology, in a passage that notes--but does not seem to rely on--an observation made by Professor Scherer in his rebuttal testimony on behalf of the shipper interests: In addition, Sinclair [via Scherer's declaration] notes that AOPL's method is a fixed-weight approach formerly used in the calculation of the Consumer Price Index but recently discarded. This change occurred because the fixed-weight approach ignored consumer substitution from high-priced goods to low-priced goods, consequently overestimating the amount of price inflation in the econo- my. Order at 61,849. But, as we said, the Commission did not actually rest its decision on this point, and that may explain why the Association's brief, and indeed that of the Commis- sion, do not address it. It is thus inappropriate for us to do so. In sum, FERC deviated from its previous methodology without any explanation responsive to the Association's objec- tions. 2. Statistical Outliers. The Association next takes issue with FERC's refusal to remove statistical outliers in conduct- ing its study. Statistical outliers are data points so extreme as to raise a question whether they may be the result of recording or measurement errors or some other anomaly. For example, some pipelines may mistakenly report their volumes in barrels, rather than barrel-miles, resulting in errors of several orders of magnitude. See, e.g., Notice of Inquiry, 65 Fed. Reg. at 47,361-62 app. A & tbl. (reporting six instances, such as a reported volume of 38 million barrels instead of 8.4 billion barrel-miles). To minimize the risk that such extreme (and erroneous) observations will bias their results, statisticians commonly use only the middle portion (e.g., the middle 50% or 80%) of the dataset for their analy- ses, Order at 61,852, or remove likely outliers in some other systematic way. In Orders Nos. 561/561-A, FERC relied significantly on Professor Kahn's 1993 analysis and defended that study's use of the middle 50% of the cost change dataset. See Order No. 561-A at 31,096; see also 1993 Kahn Study, at 9. In the current Order, however, FERC refused to adopt any similar removal of outliers. See Order at 61,852. The Association contends that FERC deviated from its previous methodology without adequate explanation. We agree. FERC first suggests that Orders Nos. 561/561-A did not rely solely on the Kahn analysis, but also rested on an analysis by Dr. Robert C. Means, which did not exclude outliers and yet came to similar results. See AOPL I, 83 F.3d at 1434. But Dr. Means used the entire dataset only after he adopted the corrections proposed by the Associa- tion's expert. See Test. of Robert C. Means, Revisions to Oil Pipeline Regulations Pursuant to the Energy Policy Act of 1992, Docket No. RM93-11-000 (Dec. 9, 1993) ("Means Study"), at 15-22; see also id. at 16, 18 (describing Professor Kahn's methodology as a reasonable response under the circumstances). And the Commission in 1993-94 stoutly de- fended Professor Kahn's removal of outliers, observing the "median is, in fact, often preferred statistically as a measure of central tendency in cases where the distribution is highly skewed." Order No. 561-A at

31,097. The Commission further cited Dr. Means in defense of Professor Kahn's method. Id. Given that insistence by the Commission, it can hardly be freed from its conventional duty to explain a change in methodology just because the results of the two 1993-94 methods--one removing outliers from an uncorrected study, and the other effecting no such removal from a corrected one--happened to coincide, especially as there is no indication that in 2000 there was any such systematic scouring of the data as that on which Dr. Means relied. Although Appendix A of FERC's Notice of Inquiry appears to correct for instanc- es where barrels had been used in lieu of barrel-miles, see Notice of Inquiry, 65 Fed. Reg. at 47,359 n.16, 47,361-62 (Appendix A), those errors represent only one class of the errors that Dr. Means had corrected in 1993, see Verified Statement of John C. Klick, Revision to Oil Pipeline Regula- tions Pursuant to the Energy Policy Act of 1992, Docket No. RM93-11-000 (Nov. 22, 1993), at 17-18 (noting modifications of operating revenues, operating expenses, net plant, and barrel-mile data). FERC's principal objection to excluding outliers was that when the dataset was narrowed from the middle 100% to 90% to 80% to 50%, the average systematically increased. Order at 61,852. FERC found these increases "troubling" and thus opted to use the entire dataset. Id. To the extent that FERC refused to exclude outliers on the ground that doing so changed the result, it obviously missed the whole point: the object of excluding outliers is to prevent extreme and spuri- ous data from biasing an analysis, i.e., affecting its result adversely. To the extent that FERC refused to adjust only because of the direction of the resulting change (upward rather than downward), refutation is (we hope) superfluous. Finally, FERC suggested that use of a complete dataset is preferable to "sampling." See Order at 61,852/2 (drawing analogy to sampling contaminated dirt). But the case for removal of outliers is independent of sampling. There was no sampling here--the dataset, however replete with errors, was complete. An objection to sampling was simply irrelevant to a claim that the errors called for removal of outliers. 3. Changes in Net Plant. The Association objects that FERC's Order estimated capital cost changes inaccurately and in an unexplained deviation from its earlier methodology. According to FERC, capital costs fall into four categories: depreciation, amortization, return on investment, and income taxes. It is undisputed that FERC's Staff study accounted for depreciation and amortization. The question revolves around return on investment and income taxes, which the Association argues should be approximated by using net plant. FERC said that net plant was an imperfect measure and might distort the analysis. Order at 61,853. Stating that the two disputed elements of capital cost are relatively minor, FERC concluded that it should not be used. Order at 61,853-54. The problem for FERC is that in Orders Nos. 561/561-A, it specifically defended the use of net plant to calculate return on investment and income taxes. Order No. 561-A at 31,098. Indeed, at the time, it was the Association that objected to net plant as inaccurate; FERC conceded the imperfections of using net plant as a proxy, but then noted that the Associa- tion offered no better solution. Id. The Commission appears to suggest that Orders Nos. 561/561-A may have used net plant to measure capital cost changes generally rather than just return on investment and income taxes. If so, then FERC's argument that deprecia- tion and amortization are a better measure than net plant would implicitly justify its dispensing with any use of net plant. Parsing both Orders Nos. 561/561-A and the 2000 Order, however, we believe that in the earlier orders FERC used net plant only for the narrow elements of investment and income taxes. First, in calculating cost changes, FERC relies primarily on data from FERC Form No. 6 submissions, which according to the 2000 Order, include data on deprecia- tion and amortization. See Order at 61,852-53 (including depreciation and amortization under "general expenses," which is a subset of the "operating expenses" listed on Form No. 6). In Orders Nos. 561/561-A, however, FERC referred to net plant as a "proxy." Order No. 561-A at 31,098. Clearly, if direct data on depreciation and amortization were already available, no proxy for them would be needed. Sec- ond, Orders Nos. 561/561-A relied heavily on the 1993 Kahn Study, which explicitly used net plant to approximate only return on investment and income taxes. See 1993 Kahn Study, at 15. Thus, having

previously used changes in net plant for one purpose despite its imperfections, FERC turned around and relied on those very imperfections to reject its use. It has offered no explanation for the change. 4. Index Adjustments. Finally, the Association chal- lenges FERC's refusal to consider two factors that it believes help undermine the case for PPI-1, factors it believes render exclusive reliance on the past six-year period inappropriate. First, it proposed an adjustment for the one-time productivity gains that the pipelines likely experienced as a result of Orders Nos. 561/561-A's change to incentive-based regula- tion. Prior to those orders, FERC regulated the industry using a rate-of-return scheme, which tended to dull the pipelines' incentive to reduce costs. Indeed, under some not implausible assumptions, pipelines had perverse incentives to "gold-plate" facilities, as all costs reasonably incurred in one period form a basis for raising rate ceilings in the next. Cf. Nat'l Rural Telecom, 988 F.2d at 177-78. The price-cap approach tends to diminish this dulling effect. Year-to-year rate increases are based on the index, and the index in turn is based on industrywide experience; so no one pipeline's cost experience has much impact on the caps to which it is subject. The Association argues that the historical cost changes ob- served in the years immediately following Orders Nos. 561/561-A were artificially depressed because of these one-time cost savings, and thus any modeling of future costs should control for them. See 2000 Kahn Study, at 23-24. The Association's second proposed adjustment was for an- ticipated future cost increases due to increased environmental and safety regulations. Among other things, the Association cites new Department of Transportation regulations imposing more elaborate employee training, increased safety testing, and design changes. AOPL Comments, at 15-16. Though they may seem distinct, the two adjustments essen- tially require FERC to perform the same task--to predict how future cost changes may deviate from the historical trend. FERC has refused to engage in such speculation, and we cannot find FERC's refusal arbitrary or capricious. Con- sistent with its congressional mandate to establish "a simpli- fied and generally applicable ratemaking methodology," EPAct, at s 1801(a), FERC opted for a purely historical analysis and has adhered to it. See Order at 61,855; Order No. 561 at 30,951 (choosing PPI-1 because it came "closest of all the indices considered ... to tracking the historical changes in the actual costs of the product pipeline industry" (emphasis added)). FERC's approach is thus unlike the one described in United States Telephone Ass'n v. FCC, 188 F.3d 521 (D.C. Cir. 1999). There the agency (the FCC), in calcu- lating the year-to-year change index for the first period covered by a rate cap, had lopped 0.5% off the historic trend line in anticipation of special productivity gains expected to flow from the switch to rate caps. Id. at 527. We thus found unexplained the agency's continuing to lop the 0.5%, as the benefits of the one-time shift could hardly be expected to go on forever. Id. As Orders Nos. 561/561-A employed no forward-looking methodology, and the agency appears to have reasonably declined to embroil itself in the complexity and iffiness of such a method, we have no basis for expecting it to adopt such a method now. * * * In summary, we conclude that FERC has neither ade- quately addressed the Association's concerns over floating- weight averaging, nor in the alternative has it articulated reasons for changing its averaging methodology. In addition, FERC failed to justify its methodological shifts regarding outliers and the use of net plant between Orders Nos. 561/561-A and the current order. We find, however, that FERC's refusal to adjust its index for past one-time produc- tivity gains and anticipated future regulatory costs is consis- tent with past practice and is reasonable. Accordingly, we remand the case to FERC for its further consideration of the first three issues, but affirm as to the last. We do not vacate, however, because it is unclear whether the remanded issues will change FERC's cost data analysis sufficiently to render the selection of PPI-1 inappro- priate. So ordered.