

Department of Health and Human Services

DEPARTMENTAL APPEALS BOARD

Appellate Division

SUBJECT: California Department of Social Services
Docket No. A-08-67
Decision No. 2285

DATE: December 4, 2009

DECISION

The California Department of Social Services (California) appealed a determination by the Administration for Children and Families (ACF) to withhold \$8,973,041 in federal funds under titles IV-B and IV-E of the Social Security Act (Act). ACF conducted a Child and Family Services Review (CFSR) of California's programs in September 2002 and found that California was not in substantial conformity with federal requirements. Subsequently, ACF found that, although California had completed the numerous action steps in its Program Improvement Plan (PIP) and met most of its target improvement goals, California had not achieved its target goals for two "statewide data indicators." Both of these indicators relate to a single performance "outcome" for children -- permanency and stability in their living situations. ACF had published these statewide data indicators (placement stability and reentry into foster care) and national standards for these indicators through a manual and memoranda issued to the states.

Based on its findings, ACF withheld funds pursuant to section 1123A of the Act. During the appeal, ACF determined that California had met an alternative improvement goal for reentry into foster care. As a result, the withholding is now based solely on ACF's finding that California did not meet its target improvement goal for placement stability, one of six items related to the single performance outcome at issue.

While California raises a number of issues on appeal, our decision here is a narrow one. Based on the record before us in this case and the reasons explained below, we conclude that the

data sets and methods used to set the standard for the placement stability indicator for California's initial review and to set California's target improvement goal for this indicator were inconsistent with the applicable regulations and with ACF's own guidance. As a result, ACF's determination that California was not in substantial conformity had no valid legal or factual basis. Accordingly, we reverse that determination.

Legal Background

The statute

Part B of title IV of the Act establishes a program for Child Welfare Services.¹ Part E of title IV establishes the Foster Care and Adoption Assistance Program. Section 1123A of the Act provides in relevant part:

(a) IN GENERAL.— The Secretary, in consultation with State agencies administering the State programs under parts B and E of title IV, shall promulgate regulations for the review of such programs to determine whether such programs are in substantial conformity with—

- (1) State plan requirements under such parts B and E,
- (2) implementing regulations promulgated by the Secretary, and
- (3) the relevant approved State plans.

Among other things, these regulations were to "specify the requirements subject to review, and the criteria to be used to measure conformity with such requirements and to determine whether there is a substantial failure to so conform;" Act § 1123A(b) (2).

With respect to "any State program found to have failed substantially to so conform," the regulations were to require the Secretary, among other things—

¹ The current version of the Social Security Act can be found at www.ssa.gov/OP Home/ssact/comp-ssa.htm. Each section of the Act on that website contains a reference to the corresponding United States Code chapter and section. Also, a cross-reference table for the Act and the United States Code can be found at 42 U.S.C.A. Ch. 7, Disp Table.

(A) to afford the State an opportunity to adopt and implement a corrective action plan, approved by the Secretary, designed to end the failure to so conform;

* * *

(C) to suspend the withholding of any Federal matching funds under this section while such a corrective action plan is in effect; and

(D) to rescind any such withholding if the failure to so conform is ended by successful completion of such a corrective action plan.

Act § 1123A(b)(4). Subsection 1123A(c) of the Act provides for appeal to the Departmental Appeals Board of a final determination that a state's program is not in substantial conformity.

The CSFR regulations

In January 2000, ACF promulgated final regulations establishing the CSFR process. 65 Fed. Reg. 4020 (Jan. 25, 2000). The regulations provide that—

ACF will determine a State's substantial conformity . . . based on the following:

- (1) Its ability to meet national standards, set by the Secretary, for statewide data indicators associated with specific outcomes for children and families;
- (2) Its ability to meet criteria related to outcomes for children and families; and
- (3) Its ability to meet criteria related to the State agency's capacity to deliver services leading to improved outcomes.

45 C.F.R. § 1355.34(a). The criteria related to outcomes include two that are related to permanency for children. Permanency Outcome 1, the only criterion at issue here, is that "children have permanency and stability in their living situations." 45 C.F.R. § 1355.34(b). The regulations also list seven systemic factors as criteria related to state agency capacity to deliver services, but none of those factors is at issue here. 45 C.F.R. § 1355.34(c).

The regulations provide that a "State's level of achievement with regard to each outcome reflects the extent to which a State" has "[m]et the national standard(s) for the statewide data indicator(s) associated with that outcome, if applicable,

. . . ." 45 C.F.R. § 1355.34(b)(2). Further, the regulations provide in relevant part:

A State will be determined to be in substantial conformity if its performance on:

- (i) Each statewide data indicator developed pursuant to paragraph (b)(4) of this section meets the national standard described in paragraph (b)(5) of this section; and
- (ii) Each outcome listed in paragraph (b)(1) of this section is rated as "substantially achieved" in 95 percent of the cases examined during the on-site review (90 percent of the cases for a State's initial review).

45 C.F.R. § 1355.34(b)(3) (emphasis added). Paragraph (b)(4) provides for the development of "statewide data indicators" for each of the described outcomes. Paragraph (b)(5) states, as relevant here:

The initial national standards for the statewide data indicators . . . will be based on the 75th percentile of all State performance for that indicator as reported in AFCARS [the Adoption and Foster Care Analysis and Reporting System] The Secretary may adjust these national standards if appropriate. The initial national standard will be set using the following data sources:

* * *

- (ii) The 1998b, 1999c, and 2000a submissions to AFCARS (or the most recent and complete report periods available) for those statewide data indicators associated with the permanency outcomes.

(Emphasis added.)

The preamble to the final rule identified seven statewide data indicators, but also stated that ACF would develop a procedures manual providing "specifics about . . . the performance indicators that will be used to measure outcomes, . . ." 65 Fed. Reg. at 4024.

The preamble also stated that the "national standard for each statewide data indicator identified above will be based on the 75th percentile of all State[s]' performance for that data indicator" 65 Fed. Reg. at 4024. In explaining why the 75th percentile was chosen, the preamble stated: "We considered using the 90th percentile and the median to establish the

national standard and rejected both because these standards, respectively, were deemed either too high or too low." Id. To illustrate this, the preamble set out a chart, showing the identified statewide data indicators and percentages representing the median, 75th percentile, and 90th percentile for each indicator, based on AFCARS data from the 1998b period (April 1-September 30, 1998). The 75th percentile value shown for the "placement stability" indicator ("% of children in care less than 12 months with no more than 2 placements") was 77%. Id.²

Under the statute and regulations, if a state does not demonstrate substantial conformity in the initial review, it must enter into a PIP. The PIP must be developed jointly by state and ACF staff and, among other things, set forth "the amount of progress the statewide data will make toward meeting the national standards." 45 C.F.R. § 1355.35(a)(1). The "methods and information used to measure progress must be sufficient to determine when and whether the State is operating in substantial conformity or has reached the negotiated standard with respect to statewide data indicators that failed to meet the national standard for that indicator." 45 C.F.R. § 1355.35(e)(1).

If ACF and a state cannot reach consensus regarding the degree of program or data improvement to be achieved, ACF "retains the final authority to assign" the degree of improvement and "will render a written rationale" for the degree assigned. 45 C.F.R. § 1355.35(d)(4). The PIP may be jointly renegotiated if, among other things, the renegotiated PIP is designed to "achieve a standard for the statewide data indicators that is acceptable to ACF." 45 C.F.R. § 1355.35(e)(4).

ACF's Procedures Manual and later issuances

In August 2000, ACF issued its Procedures Manual. Under "Permanency Outcome 1," the Manual listed seven items, including

² The preamble did not explain why the median (63%) was deemed to be too low or the 90th percentile (85%) to be too high, but the preamble did go on to say that "[w]e recognize that we have set a high standard. However, we think it is attainable and that our overall approach for moving States to the standard through continuous improvement is sound." 65 Fed. Reg. at 4025.

"Incidence of foster care re-entries (onsite and statewide aggregate data)" and "Stability of foster care placement (onsite and statewide aggregate data)." For purposes of a statewide data indicator based on AFCARS data, the placement stability "measure" was identified as the percent of children in care less than 12 months with no more than two placements.

To explain how a "national standard will be established for each of the statewide aggregate data indicators," the Manual listed the following bullet points:

- Each State and the District of Columbia submits data through AFCARS for selected time periods. (Note: The foster care standards will initially be set using 1998a and 1998b AFCARS submissions.)
- All of the data submissions for the time periods are pooled. Assuming that each of the States and the District of Columbia submitted complete data for each of two periods, there will be 102 data elements in the pool.
- All of the data submissions are then rank-ordered on a scale from highest to lowest. The point on the scale that represents the 75th percentile is the national standard for that statewide aggregate data indicator.

August 2000 Manual at 46. Like the preamble, the Manual identified the national standard for the placement stability indicator as 77%, meaning that 77% of the foster care children in care less than 12 months had no more than two placements.

The Manual did not explain why 1998a and 1998b AFCARS submissions would be used, instead of the submissions identified in the regulation.

In Information Memorandum ACYF-CB-IM-00-11 issued on December 28, 2000, ACF informed states that it would use six out of the seven statewide data indicators listed in the preamble to the final rule. With respect to the national standards for those indicators, the information memorandum stated:

We are not using the AFCARS reporting periods as described in the preamble and the regulation . . . Rather, we selected the time periods described in the following chart to establish the national standards in order to avoid using the same data to establish the standards and to determine substantial conformity on the first reviews.

ACYF-CB-IM-00-11 at 2 (emphasis added). For the placement stability indicator, the memorandum identified AFCARS data for reporting periods 1998a and 1998b. The memorandum also explained the methodologies ACF used for computing the national standards. Id. at 3.³

ACF next issued Information Memorandum ACYF-CB-IM-00-11 to correct the methodology for calculating the national standard for foster care re-entries, which is no longer at issue here.

Then, in Information Memorandum ACYF-CB-IM-01-07, issued August 16, 2001, ACF rescinded its two earlier memoranda in order to "provide updated information on the national standards" and to "provide guidance for use by States and Regional Offices in negotiating the amount of improvement necessary toward meeting the national standards through an approved" PIP. This information memorandum included the following bullet points in explaining what ACF did to "update" the national standards:

- The application of new rules that disregard skewed data points in calculating the national standards;
- The inclusion of a sampling error associated with each national standard; . . .

ACYF-CB-IM-01-07, at 2 (emphasis added).

A chart listed 86.7% as the national standard for the statewide placement stability indicator ("% of children in care less than 12 months with no more than 2 placements"). The 86.7% was determined by subtracting 1.90% for "sampling error" from the amount identified as the 75th percentile of states' performance.

The revised information memorandum also explained that ACF would "consider the sampling error for each indicator as a minimum percentage of improvement for a State to make over the course of

³ A chart listed 89% as the national standard for the placement stability indicator ("% of children in care less than 12 months with no more than 2 placements"). This figure is 12 percentage points higher than the figure given in the preamble and Procedures Manual as the 75th percentile of states' performance (77% had no more than two placements) and 4 percentage points more than the 90th percentile (85% had no more than two placements) that the preamble described as too high.

a PIP" but adjust the goal for "programmatic factors." Id. at 4.

Background of the case

The following facts are undisputed. ACF initiated the initial (first round) CFSRs in 2000. None of the 50 states passed the first round standards.

California's child welfare system is county-operated, with 58 counties. CA Ex. 1 (Declaration of Wesley A. Beers), at ¶ 17. California has the largest and most ethnically diverse child welfare population in the nation. California is also one of a small number of states that includes probation youth in its child welfare population; these youth generate some unique case movements. Id. at ¶ 29. California's first round CFSR began in February 2002, with the statewide assessment; the on-site review was performed in September 2002. Id., Attachment (Att.) 1.

ACF's final report, issued on January 10, 2003, found that California was not in substantial conformity with any of the seven outcomes set out in the regulations or with five of the seven systemic factors. ACF suspended the withholding of the \$18,244,430 ACF determined was associated with the seven outcomes and five systemic factors while California was implementing its PIP. By the end of two years, ACF rescinded the withholding of funds associated with the systemic factors. ACF gave California until September 30, 2006 to demonstrate achievement of the PIP goals for the seven outcomes. ACF ultimately rescinded the withholding of funds associated with six of the seven outcomes.

California asserts, and ACF does not deny, that by September 30, 2006 California had completed each of the 2900 tasks and each of the 141 action steps called for in its revised PIP, along with additional steps that were part of Child Welfare Reform in the State. CA Ex. 1, at ¶ 30. Specifically, California's PIP was predicated upon "two comprehensive state-wide reform efforts" -- the "California Child Welfare Services Redesign and the AB 636 California Child Welfare Outcomes and Accountability System." Id. at ¶ 24.

In a letter dated January 4, 2008, however, ACF informed California of ACF's final determination that California "did not achieve the PIP goals related to Permanency Outcome 1, Children Have Permanency and Stability in their Living Situations." CA Ex. 3, Att. 1, at 1 (emphasis omitted).

ACF found that California met four of the six goals for Permanency Outcome 1. ACF determined, however, that California did not meet the revised PIP target of 7.4% for the foster care reentry statewide data indicator or the revised PIP target of 84.4% for the placement stability statewide data indicator. The 84.4% revised target for placement stability is 1.9% more than California's revised baseline of 82.5%.⁴ ACF found, based on AFCARS data, that California had achieved 83.0% for the placement stability indicator. CA Ex. 3, Att. 1, at 1-2.

As noted above, during Board proceedings, ACF determined that California had met an alternative improvement goal for reentry into foster care. Thus, the placement stability indicator is the only item remaining in dispute.

Analysis

On appeal, California raised numerous issues. Our decision, however, is a narrow one focused on issues California raised about the data sets and methods used to set the national standard for the placement stability indicator applied to California's initial CFSR and the related target improvement goal set for California. Below, we first make findings of fact based on the record before us about the data sets and methods used. We then explain more fully why those findings lead us to conclude that, in determining that California did not achieve substantial conformity, ACF applied a standard for placement stability that was calculated using data sets and methods that were inconsistent with the regulations and with ACF's own revised information memorandum.

⁴ California's PIP as revised November 3, 2004 had as the goal for placement stability that the statewide data indicator would increase by 3.8 percentage points to 81.6% by June 30, 2005, up from a baseline of 77.8%. CA Ex. 1, Att. 9, 30. The baseline was apparently revised upward to 82.5% because some data reports were resubmitted. CA Ex. 3, Att. 1, at 2. The revised improvement goal reflects the 1.9% that ACF identified as the "sampling error" for the indicator. California points out that 10 states that were found to have achieved substantial conformity had target goals lower than California's revised baseline, and that two states had no target for placement stability in their approved PIPs. CA Br. at 29-30.

In reaching our conclusions below, we recognize that a perfect data set and methodology to evaluate performance across the diversity of state programs may not be possible. Further, we recognize that a choice of data sets involves judgment by ACF on how to balance many competing considerations, such as accuracy, administrative efficiency, and fairness. Having made choices for the initial reviews that are reflected in regulations and policy issuances, however, ACF may not properly support a determination leading to a penalty based on a standard calculated in a manner that conflicts with the wording and intent of the regulations in some respects, and with what ACF said in its own issuance in other respects. Based on what the record before us shows about the multiple and important ways ACF deviated from the regulations and its revised information memorandum with respect to the only standard at issue here, we are compelled to conclude that ACF's determination that California did not achieve substantial conformity lacks adequate legal and factual support.

A. ACF presented no evidence that the 1998a and 1999b data sets were the most recent and complete data sets available.

As noted above, the regulation required ACF to set national standards for the initial reviews using AFCARS data from specified report periods (1998b, 1999c, and 2000a) or from "the most recent and complete report periods available." 45 C.F.R. § 1355.34(b)(5). The preamble to the final rule indicated an anticipation that "AFCARS data will be derived from the 1998b, 1999c (complete Federal fiscal year) and 2000a (October 1-March 31) reporting periods," but stated that "if we have more current and complete data available, . . . we will use these data submissions to develop the standard" and that using these data submissions "will increase the number of States that participate in setting the standard." 65 Fed. Reg. at 4025.

ACF's information memorandum conceded that ACF did not use AFCARS data from the 1999c or 2000a report periods, instead using data from the 1998a report period (October 1, 1997 to March 31, 1998) in addition to the 1998b report period. California argues that ACF did not, in fact, develop a national standard at the 75th percentile of all states' performance for the placement stability indicator according to the regulation, instead using data sets that were incomplete and unrepresentative and a flawed method for calculating the 75th percentile. Among other things, California presented evidence showing that ACF not only did not use AFCARS data from the

specified report periods, but also did not use data from the most recent and complete report periods available.

AFCARS is the acronym for the data collection and reporting system required under 45 C.F.R. § 1355.40. States were required to have a system meeting the regulations and were to begin transmitting data to ACF no later than May 15, 1995. Data elements and standards are set out in appendices to 45 C.F.R. Part 1355. Federal funding was made available beginning with fiscal year 1994 for statewide automated child welfare information systems (SACWIS). 45 C.F.R. § 1355.52. One of the conditions for funding approval for a SACWIS is that it provide automated procedures and processes to meet "the Adoption and Foster Care reporting requirements through the collection, maintenance, integrity checking and electronic transmission of the data elements specified by the" AFCARS requirements. 45 C.F.R. § 1355.52(b). States are required to transmit reports semi-annually "within 45 days of the end of the reporting period (i.e., by May 15 and November 14)." 45 C.F.R. § 1355.40(b). Thus, at the time ACF issued its updated national standard for the placement stability indicator in August 2001, states were to have made submissions for reporting periods through the first half of FY 2001.

Despite California's challenge to ACF's use of the 1998 data, ACF does not assert in its briefs that the 1998 data sets were the most recent and complete data sets available, nor provide any evidence to show that they were. The evidence California submitted indicates to the contrary that, with each additional year of AFCARS reporting, the reports were more complete. Specifically, California submitted reports from the HHS Office of Inspector General (OIG) and the Government Accountability Office (GAO). CA Exs. 6 and 7.

The OIG report, issued in March 2003, directly addressed the issue of completeness of AFCARS data. The OIG found, based on the annual AFCARS report, that, for FY 1998, "no data were reported for 10 states and selected data were missing for 8 additional states." CA Ex. 7, at 5. For FY 1999, however, only two states reported no data, with 10 states missing some data from their profiles. Id. In other words, six more states submitted complete data in FY 1999 than in FY 1998.

The GAO report, issued in July 2003, discussed the challenges states were facing in developing and implementing SACWIS systems and in reporting child welfare data. CA Ex. 6, Highlights page, see also at 24. In response to GAO's draft report, ACF conceded

the data problems, but took issue with GAO's characterization of ACF's efforts to improve data as "recent," instead describing them as "ongoing" and listing ACF issuances from 1994 to 2003 to provide technical and other information to states. Id. at 51 to 54.⁵ In the absence of any evidence from ACF to the contrary, we infer that the data reporting was improving after 1998, not getting worse, since states were implementing SACWIS systems in this period, and, beginning in 1994, ACF had been providing assistance, making efforts to improve data, and issuing technical and other information.

Thus, given that more states submitted complete data in FY 1999 than in FY 1998 and that other evidence indicates that the reporting was improving, we find that the 1998a AFCARS data set was not more complete than the data sets from the more recent reporting periods specified in the regulations, 1999c and 2000a. The deviation from the regulatory provisions in this regard therefore introduced a greater potential for error in comparing California's performance with other states' performance.

B. Methodologies used to compute the national standard for placement stability and the target goal for improvement were not consistent with ACF's revised information memorandum or with statistical principles.

California presented evidence raising concerns about the validity of ACF's methods and calculations in determining a national standard for placement stability and California's target goal for improvement. The evidence was presented in the initial declaration by John R. Schuerman, Ph.D. CA Ex. 5. Much of what Dr. Schuerman says in his declaration about the computations is shown in Attachment 2 to the declaration. Dr. Schuerman attested that Attachment 2 contains information obtained from ACF worksheets but presented in a more readable version. CA Ex. 5, at ¶ 26. ACF does not deny that the

⁵ We note that ACF's website says about AFCARS data: "Over the years, the accuracy and completeness of the data reported to AFCARS have improved dramatically. Therefore, care should be taken in interpreting trends appearing in the earlier years because they may be primarily a function of data quality improvement." See http://www.acf.hhs.gov/programs/cb/stats_research/afcars/tar/report12.pdf.

information in Attachment 2 is from ACF worksheets, nor does ACF contend that the more readable version is inaccurate in any respect.

Based on the information in Dr. Schuerman's declaration and Attachment 2 to that declaration, we make the following findings of fact, which we set out in numbered bold sentences, followed by a discussion of the evidence.

1. ACF omitted data points beyond those that the revised information memorandum explained would be deleted.

Information Memorandum ACYF-CB-01-07 provided a rationale for the revised national standards set out in that memorandum. That memorandum stated:

When recalculating the national standards, we applied new rules that disregarded extreme data points in individual States' data that were used to calculate the national standards. We excluded State data indicators that were 0% and 100% from the data points used to determine the standards, believing that such extreme values were unlikely to be accurate and would skew the standards in one direction or another.

Id. (emphasis added). In fact, however, it is undisputed (and the information obtained from ACF shows) that other data points were deleted which were neither 0% nor 100%, but were all at the low end of the scale. CA Ex. 5, Att. 2. As explained in detail below, California's expert attested, without rebuttal, that these deletions had the effect of skewing the data and setting a higher standard. In other words, the "updated" standards were not calculated in a manner consistent either with what ACF said it was doing (deleting only data points that were 0% or 100%) or with the stated intent - to avoid skewing the data.

The information obtained from ACF shows that ACF dropped the additional numbers as part of a process to make the data "fit" a "normal distribution." CA Ex. 5, Att. 2. The data were not obtained by drawing a statistical sample randomly from a universe, however, but were drawn from AFCARS reports for those states who reported the relevant data. Dr. Schuerman attested:

The justification for this fitting of the data to a normal distribution has not been given by ACF. The normal curve is a theoretical statistical distribution. Very few quantities in nature are normally distributed. Normal statistical practice

would have been to simply determine the 75th percentile of the distribution as it stood rather than first forcing the data into a normal distribution.

CA Ex. 5, at ¶ 21. ACF presented no evidence to rebut this assertion. Nor did ACF present any evidence to rebut Dr. Schuerman's description of the computer program routine used to identify data points (or "observations") that did not "fit" a normal distribution and therefore should be "dropped" from the data set as being likely errors. Id., at ¶ 22. According to Dr. Schuerman, the computer program routine is "a quite mechanical process which ignores substantive considerations, not taking into account whether the observations make sense in the real world." Id.

ACF has provided no rationale to us for deleting further data points beyond the 0% and 100% figures it said it did not think were accurate. Moreover, as discussed below, deleting additional data points made the data even more incomplete than they already were and less representative - which is inconsistent with the regulatory goal of using the most complete data sets to include as many states as possible in setting a national standard.

2. The 1998 data sets ACF used in setting the standard for placement stability were already incomplete and not representative before ACF omitted additional data points.

There were potentially 104 data points for the 1998a and 1998b report periods (two data points for each of 52 jurisdictions -- the 50 states, the District of Columbia, and Puerto Rico). CA Ex. 5, at ¶ 23. For the placement stability indicator, 38 states and the District of Columbia submitted data for both time periods, so there were 78 data points in the original data set. Id., Att. 2. Data from the following 12 states were missing: Alaska, Kentucky, Florida, Massachusetts, Michigan, Nebraska, Nevada, New Hampshire, New York, Ohio, South Dakota, and Tennessee; data from Puerto Rico were also missing. Id. Thus, "approximately 24% of the jurisdictions were not represented in the computation of the national standards representing about 30% of the child welfare population in the U.S." Id., at ¶ 24 (footnote omitted).

Dr. Schuerman attested that the "39 jurisdictions that were represented cannot be considered an unbiased sample of all states," particularly since some large population states such as Florida, Michigan, Massachusetts, and New York, were not

included. Id. This is significant because the calculations used by ACF treated data from all of the states the same, without weighting by population or child welfare caseload size. Dr. Schuerman attested that this treatment "tends to give inordinate weight to smaller states and does not capture the state of child welfare practice in the United States as a whole." CA Ex. 5, at ¶ 8. Dr. Schuerman also attested that another difficulty with the incompleteness of the data arises because of differences among the states in how they reported placement data. CA Ex. 5, at ¶ 9. This testimony is supported by the OIG and GAO reports identified above, which discuss differences in how states reported placement data, including for juvenile justice populations. CA Ex. 6, at 24, 69; CA Ex. 7, at ii.

ACF did not deny Dr. Schuerman's assertions that the 1998 data sets were incomplete and not representative or present any evidence to rebut those assertions.

3. ACF did not actually used the 75th percentile of all states' performance.

ACF's Procedures Manual said that all of the data submissions would be "rank-ordered on a scale from highest to lowest" and that the "point on the scale that represents the 75th percentile is the national standard" for a statewide data indicator. August 2000 Manual at 46. As the information obtained from ACF shows, this is not what was done in setting the standard for the placement stability indicator, however. Instead, as mentioned above, the data were first manipulated so that they would "fit" to a "normal distribution" (or "normal curve" as shown on a graph). This resulted in dropping some data points before the 75th percentile of the remaining data points was determined. CA Ex. 5, Att. 2. The computer routine used was an "iterative process" which may involve several steps in which observations are dropped and then the resulting distribution checked to see if it has become close enough to a normal distribution. Id., at n. 24.

The information obtained from ACF worksheets shows that the 78 observations in the original data sets from 1998 ranged from 20.96% to 99.08%. Id., Att. 2. After these data had been subjected to the computer program routine to determine normality (which took three "iterations"), six out of the 78 observations had been dropped, leaving only 72 data points. Id. No observations from the top of the distribution were dropped. Id.

The information from ACF shows, as Dr. Schuerman asserted, that this "dropping of observations only at the low end upwardly biased the determination of the 75th percentile, resulting in a higher national standard." CA Ex. 5, at ¶ 25. This concern is particularly acute, given the evidence that the 1998 data sets were already incomplete and unrepresentative.

4. ACF did not correctly calculate the amount treated as "sampling error" and used to set the standard and the target improvement goal for the placement stability indicator.

ACF said in its revised information memorandum that it would calculate the "sampling error" associated with its determination of the 75th percentile, reduce the standard by that amount, and use the "sampling error" to set a minimum improvement goal. Specifically, the memorandum stated:

In calculating the new national standards, we included the sampling error associated with each data indicator as part of the actual standard. Even with the use of more accurate data to calculate the national standards, we realize that there is a certain range within which a State's data can be considered to meet the standard statistically. We chose to incorporate the sampling error into the standards because we do not believe it is appropriate to subject a State to a potential penalty if the State is within a statistically acceptable range of the national standard.

ACYF-CB-IM-01-07, at 2.

The unrebutted evidence in the record shows that ACF did not, in fact, correctly take sampling error into account.

Dr. Schuerman attested, and ACF does not deny, that the calculations used to implement ACF's decision to take sampling error into account involved calculating the "standard error of the mean of the [72] values" left after the procedure to fit the data to a normal distribution, then constructing a "single sided 95% confidence interval," and subtracting the size of this interval from the 75th percentile to set the national standard. CA Ex. 5, at ¶ 28.

This calculation assumed that the sampling error of the 75th percentile is the same as the sampling error of the mean (or average) value. Citing to a treatise on statistical sampling, however, Dr. Schuerman attested that--

the standard error of the 75th percentile is considerably larger than that of the mean, over one-third larger in normal distributions. Thus, using the standard error of the mean underestimated the adjustment to the 75th percentile by over one-third.

Id. (footnote omitted). ACF presented no evidence to rebut this assertion.

Nor did ACF present any evidence to rebut Dr. Schuerman's opinion that the "sampling error" calculation was also flawed because it failed to take into account the fact that the data points used to calculate the national standard were not independent. Dr. Schuerman attested that independence of observations is a requirement that arises out of the probability theory from which inferential statistical procedures are derived. Id. at ¶ 29, and n. 34 (citing a statistics text). Dr. Schuerman pointed out, and the information from ACF shows, that the 72 data points used to calculate the standard for placement stability included two data points from each of 35 states and one each from two states. According to Dr. Schuerman, the "data from the 35 states with two observations are not independent, that is, a State's 1998A observation is likely closely related to its 1998B observation." Id.

Dr. Schuerman attested that he applied a test that measures the association (or correlation) between two values for a set of units, and found that the "correlation between the [1998]A and [1998]B values for the entire set of 78 observations" is "0.98, which indicates a high degree of association." Id. at ¶ 29 and n. 35.⁶ According to Dr. Schuerman's unrebutted opinion, this correlation "affects the validity of the standard error of the mean" since the "calculation of this value involves dividing by the square root of the number of independent values used." Id. (emphasis added). ACF's divisor was the square root of 72, the number of observations left after ACF had dropped some numbers. Dr. Schuerman attested, however, that if the correct number of independent values of the data set were used, "the divisor should have been close to the square root of 37." Id.

⁶ Under this test for independence, a perfect positive association is indicated by a "correlation coefficient" of +1, with 0 indicating no association. Id.

This is important because dividing by a much smaller number would have resulted in a larger figure for sampling error and a larger downward adjustment to the 75th percentile, and, therefore, a lower standard than the one ACF used here.

5. Dr. Schuerman's testimony and opinions are reliable.

In making the above findings, we specifically determine that the attestations of California's expert supporting these findings are reliable. ACF presented no evidence or opinion whatsoever to contradict Dr. Schuerman's assertions on these matters, but argues that California's evidence is simply the "isolated opinion of a single individual." ACF Br. at 36. We reject this argument because:

- o Dr. Schuerman is highly qualified, as shown by his attestation and curriculum vitae, which ACF does not dispute. CA Ex. 5, at 1-2, and Att. 1. He is Emeritus Professor of the School of Social Service Administration and Faculty Associate at the Chapin Hall Center for Children at the University of Chicago. He has a Ph.D in social services from the University of Chicago and has taught statistics and research methods for several decades to graduate students in social services and public policy at the University of Chicago. He has written textbooks on multivariate statistical analysis, was a senior co-author of a book evaluating family preservation programs in Illinois, and has published many relevant articles. His work has involved a number of studies for the federal government. He also has been the editor of *Social Service Review*, the leading scholarly journal in the field.
- o His assertions of fact about the data sets and the calculations are based on and consistent with the information he obtained from ACF that is included as Attachment 2 to his declaration.
- o He supported his key statements about statistical sampling principles and methods with citations to textbooks on statistics.
- o With its reply brief, California submitted the declarations of two other highly qualified experts, who concurred in Dr. Schuerman's opinions. Specifically, California presented declarations from Mark Courtney, Ph.D., Professor and Ballmer Chair for Child Well-Being in the School of Social Work at the University of Washington, and from Charles L.

Usher, Ph.D., the Wallace H. Kuralt, Sr. Professor of Public Welfare Policy and Administration in the School of Social Work at the University of North Carolina at Chapel Hill. CA Supp. Exs. 4, 5.

ACF had an opportunity to respond to Dr. Schuerman's declaration, as well an opportunity for a surreply to the additional evidence California presented with its reply brief. While ACF presented some evidence to rebut the declarants' opinions about ACF's use of the statewide placement stability indicator, ACF presented no evidence whatsoever to rebut the evidence California presented on the flaws in computing the national standard (and associated "sampling error") for that indicator.⁷ Nor did ACF seek to cross-examine any of California's declarants. Thus, we find California's evidence regarding the computations to be reliable.

C. ACF did not follow the applicable requirements to establish (or appropriately adjust) the initial national standard for placement stability applied to California.

The regulations provide for the determination of whether a state is in substantial conformity to be based, in part, on statewide data indicators and national standards developed in accordance with the regulations. The regulations provided for national standards for the initial CFSRs to be based on the 75th percentile of all states' performance on statewide data indicators, as determined using AFCARS data sets from 1998b,

⁷ ACF's surreply brief does cite to the preamble to the final CFSR regulations and to ACF Exhibits 6 and 7 to support the proposition that ACF was not required to use the "best science available" and that ACF reached reasoned decisions after receiving and weighing comments. ACF Surreply at 36-37. ACF Exhibit 7 identifies members of an ACF "Consultation Group," and ACF Exhibit 6 is a 1998 report on how "outcome measures" were developed. Exhibit 6, at 7, identifies a measure for placement stability but does not discuss how to calculate a national standard for that measure. Moreover, Dr. Schuerman's testimony at issue here is not a challenge to the regulations, as ACF's surreply assumes, but to what was done in 2001 in calculating the updated national standard for placement stability for the initial reviews. ACF does not allege that the Consultation Group was aware of and approved those calculations.

1999c, and 2000a or the most recent and complete data sets available.

ACF acknowledges in its brief that it is bound by the regulations (as we are). The regulations reserve some discretion to ACF in establishing or adjusting national standards. The regulation and its preamble, however, limited the discretion ACF had in establishing the initial national standards and in adjusting the standards.

With respect to establishing the initial standards, ACF had the option to use data from the AFCARS reporting periods described in the regulations or data from "the most recent and complete report periods available." 45 C.F.R. § 1355.34(b)(5). The stated intent of the regulations in allowing for the alternative use of a data set other than the ones specified in the regulation in determining the 75th percentile of "all states' performance" was to "increase[e] the number of States that participate in setting the standard." 65 Fed. Reg. at 4024 - 4025.

ACF expressly declined to use "the AFCARS reporting periods as described in the preamble and regulations." See, e.g., ACYF-CB-IM-00-11. ACF in fact used less recent data from the 1998a reporting period, rather than data from the 1999c and 2000a reporting periods. As we find above, however, the data sets used were not more complete than the later data sets (and, indeed, were incomplete in important respects).

Information Memorandum ACYF-CB-IM-00-11 says ACF used only 1998 data because ACF "had learned since the publication of the final rule, that it would be an inappropriate methodology to use the same data to develop the standards that would be used later to measure a State's performance." No explanation was given then (or in briefing to us) about why using the specified data sets would be inappropriate or about why ACF could not use 1999c or 2000a data for initial reviews of periods starting after those reporting periods had ended. While the determination not to use the data sets described in the regulations might represent a reasonable policy choice in isolation, ACF's reliance on the older and very incomplete 1998a and 1998b data sets conflicts with the policy stated in the regulations and preamble to use the most recent and complete data to increase state participation.

We recognize, as stated above, that section 1355.34(b)(5) provides for adjustments to the national standards for statewide

data indicators. ACF's briefs do not, however, rely on the adjustment provision in the regulations, and we conclude that it does not apply here, for the following reasons:

- o Section 1355.34(b)(5) states that the "Secretary may adjust these standards if appropriate" (i.e., the standards described in the regulations). ACF did not, however, simply adjust the regulatory standards. Instead, it established different standards based on different data sets that were not the most recent and complete sets and then adjusted those standards.
- o The preamble to the final CFSR rule explained the authority that was being retained as follows:

The standards will not change every year. Rather, we have retained our authority to periodically review and revise the standards if experience with the reviews indicates adjustments are necessary.

65 Fed. Reg. at 4025 (emphasis added). ACF does not claim that its choice about what data to use was based on its experience with the reviews.

- o As discussed above, the adjustments in fact made to the standard for the placement stability indicator were not the adjustments ACF's revised information memorandum indicated it had determined were appropriate. ACF did not follow its own issuance and delete only data points of 0% or 100%. Nor did ACF correctly implement its policy choice to use "sampling error" to reduce the standard and to set improvement goals. Instead, as we found above, the calculations used the sampling error of the mean of the data used, rather than the sampling error of the 75th percentile of the data used. Nor did the calculations take into account that the data points used were not independent. These are not merely minor or procedural flaws, but were flaws that were highly prejudicial to California.

In sum, the adjustments were neither in accordance with the regulations nor in accordance with ACF's own policy.

Conclusion

For the reasons stated above, we reverse ACF's determination that California did not achieve substantial conformity and reverse the withholding of funds based on that determination.

_____/s/
Leslie A. Sussan

_____/s/
Constance B. Tobias

_____/s/
Judith A. Ballard
Presiding Board Member