2014-2015 Nationwide Adult Medicaid Consumer Assessment of Healthcare Providers and Systems (NAM CAHPS) Nonresponse Bias Analysis

Executive Summary

In 2014 and 2015, a nationwide survey of Medicaid beneficiaries was conducted by NORC at the University of Chicago for the Centers for Medicare & Medicaid Services (CMS). This survey sought to understand beneficiaries' experience of care using the Consumer Assessment of Healthcare Providers and Systems (CAHPS) survey. The sample design was created to be broadly representative of Medicaid beneficiaries, however, there were certain criteria applied that excluded portions of the broader Medicaid population. In particular, the sample design created four unique, hierarchical beneficiary groupings:

1. Full Dual - adults dually eligible for Medicaid and Medicare (Duals);
2. Disabled - adults (non-duals) with disabilities based on program eligibility criteria (Persons with Disabilities);
3. Managed Care - adults (non-duals, non-disabled) enrolled in a managed care organization (Managed Care, or MC); and,
4. FFS PCCM - adults (non-duals, non-disabled) enrolled in a fee-for-service (FFS) provider or who were enrolled in a primary care case management plan (PCCM).

This report presents analysis of differences between the 2014-2015 Nationwide Adult Medicaid CAHPS (NAM CAHPS) survey respondents, the targeted population for this survey, and the Medicaid population in general. Some key highlights found in the report are:

- The weighted sample of respondents mimicked the set of eligible beneficiaries for most variables, though there were some differences:
    - Beneficiaries in the Full Duals and Disabled groups were more likely to respond than those in the Managed Care or Fee-For-Service strata;
    - Older beneficiaries were more likely to respond than younger beneficiaries; and,
    - Beneficiaries living in rural areas were more likely to respond than those in urban areas.
- In multivariable modeling, age and beneficiary grouping were the strongest predictors of whether a beneficiary would respond.
- The weighted sample of respondents differed from external population benchmarks in the following ways:
    - Higher percent of 45-64 year olds and fewer 18-20 year olds;
    - Higher percent of Non-Hispanic Whites and fewer Non-Hispanic Blacks and Hispanics;
    - Higher percent of females; and,

- Higher proportion from the Northeast and lower proportion from the South.

This report is divided into five sections. Section 1 gives an overview of the survey project, including the methodology implemented during sample design, sample selection, and weighting. Section 2 analyzes distributional differences between the targeted population of eligible beneficiaries and the final sample of completed surveys. Section 3 delves further into this analysis through logistic regression and odds ratios. Section 4 compares the final weighted sample to known external benchmarks. Section 5 presents a brief discussion of key takeaways from this report.

1.0 Overview of Background and Sample Design

To support states in collecting and reporting CAHPS data, the Centers for Medicare & Medicaid Services' (CMS) Center for Medicaid and CHIP Services (CMCS) contracted NORC at the University of Chicago (NORC) to conduct the first-ever Nationwide Adult Medicaid Consumer Assessment of Healthcare Providers and Systems (NAM CAHPS) survey.[1] The goal of the survey was to obtain national and state-by-state estimates of adult Medicaid enrollees' experience of care, including access, utilization, and satisfaction with care across different financing and delivery models (e.g., managed care and fee-for-service (FFS)) and population groups (e.g., dually eligible individuals and individuals with disabilities who are not dually eligible). While the overall response rate was 23.6%, there were varying levels of response by key strata used in the design of the project (see Table 1 below). This is comparable to other health care related surveys. The California Health Interview Survey (CHIS) has a response rate of 21.5% for completion of the screener, though this survey is telephone based.[2] A broader study focused on CAHPS adult Medicaid and commercial data submitted to the National Committee for Quality Assurance (NCQA) found that commercial plans had a response rate of 48% while Medicaid plans had a response rate of 38%.[3]

This report provides results from nonresponse bias analysis to understand where potential nonresponse bias may exist within the 2014-2015 NAM CAHPS survey.

Table 1.1: Response Rates by Strata, Overall for the 2014-2015 NAM CAHPS Survey.

Stratum	Completes	Partials	Eligible Sampled Cases	Response Rate
Overall	272,679	1,045	1,159,768	23.6%
Full Dual	91,456	341	244,189	37.6%
Disabled	76,704	266	249,210	30.9%
Managed Care	57,673	232	374,838	15.4%
FFS-PCCM	46,846	206	291,531	16.1%

[1] The Agency for Health Care Research and Quality refers to the survey as the Nationwide Adult Medicaid Survey (NAMS).

[2] California Health Interview Survey. CHIS 2015-2016 Methodology Series: Report 4 – Response Rates. Los Angeles, CA: UCLA Center for Health Policy Research, 2017.

[3] Weech-Maldonado, Robert, et al. "Survey response style and differential use of CAHPS rating scales by Hispanics." *Medical care* 46.9 (2008): 963.

The target population for the NAM CAHPS survey was adults ages 18 and older as of December 31, 2013 who were enrolled in Medicaid for each month of the first quarter of federal fiscal year 2014 (FFY 2014 Q1, October 2013 – December 2013), enrolled during the month prior to the start of sampling (August 2014)[4], and not residing in an institutional setting.[5] It should be noted that there were specific exclusions for the target population. During the planning phase of the project, it was decided the following categories would also be excluded from the set of eligible beneficiaries:

- Partial-Duals – beneficiaries who qualify to have Medicaid pay some of the expenses they incur under Medicare;
- Family Planning Waivers – beneficiaries with coverage limited to family-planning related services;
- Unknown Managed Care Plan – beneficiaries with an unknown managed care plan; and,
- Unknown contact information – beneficiaries with no given mailing address or phone number.

Outside of the exclusions noted above, the sample was designed to capture four key sub-groups of adult Medicaid enrollees. The main stratifiers were the states (including the District of Columbia) and four mutually exclusive enrollee groupings based on program eligibility. Group classification (strata) was determined using the hierarchy shown below.

- Full Dual - adults dually eligible for Medicaid and Medicare (Duals);
- Disabled - adults (non-duals) with disabilities based on program eligibility criteria (Persons with Disabilities);
- MCO - adults (non-duals, non-disabled) enrolled in a managed care organization (MCO); and,
- FFS-PCCM - adults (non-duals, non-disabled) enrolled in a FFS provider or who were enrolled in a primary care case management plan (PCCM).

Other than the stratum fields of each state and enrollee groups, no additional explicit stratifying variables were used in the sample design. However, additional fields were used as control variables (i.e., implicit stratifiers[6]) in the sample selection process. These were variables that were readily available for eligible beneficiaries, and therefore could be leveraged to create representative samples. The set of eligible beneficiaries and associated variables is defined as the sampling frame throughout the remainder of the report.

---

[4] Due to the time-lag in access to final CMS and state beneficiary data files, it was decided that having a more recent enrollment time period included would be beneficial in removing persons no longer eligible for the survey. The sample selection process for the survey began in September 2014.

[5] Defined as having no long-term care claims during calendar year 2013.

[6] Implicit stratifiers make sure the sample is representative of the population relative to key characteristics. Unlike explicit stratifiers, implicit stratifiers may not provide sufficient sample sizes to allow detailed analysis for those stratifiers.

We performed the following procedures to select the sample size. Within each of the four strata within each state, we sorted the sampling frame records using a hierarchic serpentine sort[7] with the following control variables (in the order presented):

- SEX - three possible groups: Male, Female, and Unknown;
- ZIP – ZIP code (include +4 when available), missing ZIPs are coded as Unknown, and form an additional sort category;
- DOB – the enrollee's date of birth.

Race and ethnicity was considered as part of the hierarchical sorting. After reviewing the potential sampling frame data, it was found that a large portion of the race and ethnicity variables were missing from the state-supplied data (see table 2.3 below). Additionally, it is likely that different protocols were implemented across states for filling in these variables (e.g., self-identified race and ethnicity versus administrative identified), mitigating the potential utility during sampling. Because of the incomplete and potentially inconsistent nature of the race and ethnicity variables on the sampling frame, 5-digit zip codes were used as a proxy for race and ethnicity, which assumed that there was some geographic clustering of minorities.[8] This was specifically used as a proxy for race and ethnicity, but there may be other underlying socio-demographics (e.g., socio-economic status) that are also geographically clustered that would have benefited from this sample design.

The overarching goal of the sample design was to attain 1,667 completed surveys in each stratum within each state. Given anticipated differential response rates based on the pilot study that was fielded in early 2014, each stratum had a specific targeted number of sampled cases: Full-Duals and Disabled sampled 5,556 cases, FFS-PCCM targeted 7,407 cases, and MCO targeted 11,110 cases. After sorting the file, we used a systematic sample selection with a random start point to select the number of enrollee records within each state and stratum. If a stratum within a state did not have the targeted number of cases, then a census was taken of all cases available. Also, some states did not have FFS-PCCM or MCO; therefore, these strata were not used for those states. During the sample design process, four states were unable to participate and were excluded from the NAM CAHPS survey (i.e., Alaska, New Hampshire, North Dakota, and Wisconsin).

At the end of the data collection process, NORC created weights to account for survey nonresponse. Base weights were first created to reflect the probability of selection within each of the strata within a state. A general nonresponse weighting adjustment was then created to account for survey nonresponse. These adjustments were done within state, stratum, age group, and sex. The final survey weights were controlled to these variables based on the total number of beneficiaries within each group identified during the construction of the sampling frame. The

---

[7] In hierarchic serpentine sorting, one sorts by the first control variable (implicit stratifier) in ascending order. Then, within the first level of the first control variable, the procedure sorts by the second control variable in ascending order. Within the second level of the first control variable, the procedure sorts by the second control variable in descending order. Sorting by the second control variable continues to alternate between ascending and descending sorting throughout all levels of the first control variable. The alternating of sort order within levels of the prior control variable continues for all remaining control variables. See: Chromy, J. R. (1979), "Sequential Sample Selection Methods," ***Proceedings of the American Statistical Association, Survey Research Methods Section***, 401–406.

[8] Charles, Camille Zubrinsky. "The dynamics of racial residential segregation." *Annual review of sociology* 29.1 (2003): 167-207.

data were not post-stratified to external population control totals because there were no known sources of external data that would match the targeted population of interest for the 2014-2015 NAM CAHPS survey based on the criteria identified above. Therefore, the only potential source of population totals were aggregated sampling frame information.

While our sample was selected in such a way as to be as representative as possible of those beneficiaries that were included in the sampling frame, not everyone selected for the survey responded. The weighting process described above assumed that those who respond to a survey are similar to and represent those who do not respond to the survey, but this discounts the potential for bias associated with "self-selection". In other words, there may be reasons an individual chooses to respond (or not respond) to a survey that may be associated and/or correlated with outcomes from the survey. This raises the risk that those who respond have differing answers to survey questions than those that do not respond, which the weighting process is unable to account for. The lower the overall response rate to the survey, the higher the likelihood for this kind of bias, commonly termed *nonresponse bias*.

The analyses presented below attempt to assess the likelihood for nonresponse bias using some standard techniques within the industry, which is by:

- Comparing respondents to nonrespondents: Section 2 compares distributions of respondents to those in the selected sample based on characteristics known for both respondents and nonrespondents. Section 3 creates a similar analysis using logistic regression to tease out characteristics significantly associated with the propensity to respond.

- Comparing respondents to the population: Section 4 shows the results of comparing the survey respondents to known benchmarks While it was noted that there were no exact external benchmarks that met the specific criteria used for the 2014-2015 NAM CAHPS survey, it is still helpful to compare to the general Medicaid population to understand what differences may exist between the two populations.

The methods above try to assess whether those who responded to the survey would have different survey responses than those who did not respond, as well as assessing differences compared to the general Medicaid population. If there are differences, identifying what characteristics available for both respondents and nonrespondents might be associated with survey responses can help inform differences and whether the weighting procedures mitigated the potential for nonresponse bias in the final weighted results.

2.0 Population Distributions

The initial nonresponse analysis consisted of reviewing distributions of the target population versus the sample. As noted above, there were specific criteria implemented in constructing the sampling frame of eligible beneficiaries for this survey, and the sample was selected in such a way that it was representative of those eligible for the 2014-2015 NAM CAHPS survey. Therefore it is important to show the comparison between the selected sample and the respondents to determine whether the respondents are representative of the targeted population.

This portion of the analysis shows the distribution of the selected sample and respondents for key socio-demographic variables that were available for all beneficiaries on the sampling frame. The tables in this section present this initial analysis. The "Sample Percent" column shows the distribution of sampled cases across the categories, while the "Completes Percent" shows the distribution of respondents. The final column shows the difference between the two columns. In cases where the difference is greater than zero, this indicates more beneficiaries responded in that category than expected. Where the difference is less than zero, this indicates less beneficiaries responded in that category than expected. Because we are comparing respondents versus the full sample which included nonrespondents, only variables available on the sampling frame were used for this initial assessment. This was to understand the differential response rates based on how the sample was selected from the sampling frame.

Table 2.1 below presents the sample and respondent distributions by strata. As noted in Section 1 above, the MCO and FFS-PCCM were sampled at higher rates than the Full Duals and Disabled strata, which is reflected in the higher percentage of sampled cases for those two categories. However, fewer cases from these two strata responded to the survey. This is reflected in both the distribution of completed cases ("Completes Percent"), but more importantly in the large differences presented in the final column. While there are larger differences between respondents and nonrespondents by strata, some of this is likely confounded with age differences that are inherently present in these group. The Full Duals and Disabled tend to be older beneficiaries, who were more likely to respond to this survey.

Table 2.1: Distribution of Sample and Completes by Strata for the 2014-2015 NAM CAHPS Survey.

Stratum	Sample Percent*	Completes Percent*	Difference
Full Dual	21.5%	33.5%	12.0%
Disabled	21.5%	28.1%	6.7%
Managed Care	32.0%	21.2%	-10.9%
FFS-PCCM	25.0%	17.2%	-7.8%

*Detail may not sum to total due to rounding.

Table 2.2 shows the distributions of sampled beneficiaries and respondents by age group. As the table shows, those 45 years of age and older are more likely to respond.

Table 2.2: Distribution of Sample and Completes by Age Group for the 2014-2015 NAM CAHPS Survey.

Stratum	Sample Percent*	Completes Percent*	Difference
1: 18-44	60.4%	39.3%	-21.1%
2: 45-64	27.2%	40.8%	13.7%
3: 65-74	6.7%	11.3%	4.6%
4: 75+	5.7%	8.6%	2.9%

*Detail may not sum to total due to rounding.

The distribution of sampled cases versus respondents by race and ethnicity is presented in Table 2.3 and shows limited differences between the sampled cases and respondents. While we see a slightly higher percent of Non-Hispanic White respondents (and slightly lower percent of Non-Hispanic Black and Hispanic respondents), these differences are fairly small in comparison to those seen in Tables 2.1 and 2.2. It should be noted that the race and ethnicity assignment used for Table 2.3 was based on what was reported in the CMS and state data files rather than what was reported by survey respondents; therefore, a large number of values are missing. The race and ethnicity assignment was based on what was reported in the sampling frame, as opposed to self-reporting from the survey. Since we do not know how nonrespondents would self-identify their race and ethnicity, Table 2.3 presents the response rates based on the sampling frame information available for the full sample that was selected. Additionally, how a respondent self-reported their race and ethnicity did not always align with what was reported in the sampling frame, especially given the large amount of Unknown/Missing reported in the sampling frame data. (Table 4.2 presents information around the respondents' self-reported race and ethnicity.)

Table 2.3: Distribution of Sample and Completes by Race and Ethnicity for the 2014-2015 NAM CAHPS Survey.

Stratum	Sample Percent*	Completes Percent*	Difference
1 – Non-Hispanic, White	28.8%	30.3%	1.6%
2 - Non-Hispanic, Black	16.0%	14.4%	-1.6%
3 - Hispanic	5.3%	4.0%	-1.4%
4 - Non-Hispanic, Other**	3.6%	3.6%	0.1%
5 - Unknown/Missing	46.3%	47.7%	1.4%

*Detail may not sum to total due to rounding.
**Non-Hispanic, Other includes Asian/Pacific Islander, American Indian/Alaska Native, as well as those with multiple races.

Table 2.4 shows the distributions of sampled beneficiaries and respondents by sex. There does not appear to be any noticeable difference in response by sex.

Table 2.4: Distribution of Sample and Completes by Sex for the 2014-2015 NAM CAHPS Survey.

Stratum	Sample Percent*	Completes Percent*	Difference
Female	66.2%	66.4%	0.2%
Male	33.8%	33.6%	-0.2%

*Detail may not sum to total due to rounding.

Table 2.5 shows the distributions of sampled beneficiaries and respondents by the four census regions. There does not appear to be any noticeable difference in response across the four regions.

Table 2.5: Distribution of Sample and Completes by Region for the 2014-2015 NAM CAHPS Survey.

Stratum	Sample Percent*	Completes Percent*	Difference
Northeast	16.6%	15.9%	-0.7%
Midwest	22.2%	23.4%	1.2%
South	35.6%	34.9%	-0.8%
West	25.6%	25.9%	0.3%

*Detail may not sum to total due to rounding.

Table 2.6 shows the distributions of sampled beneficiaries and respondents by urbanicity based on grouping the Primary RUCA codes into urban and rural areas (Hall et al 2006)[9]. Note that a vast majority of the eligible beneficiaries live in urban areas. While there are some differences in urbanicity between the sample and respondents, this does not appear to have an impactful change in the overall distribution.

Table 2.6: Distribution of Sample and Completes by Urbanicity for the 2014-2015 NAM CAHPS Survey.

Stratum	Sample Percent*	Completes Percent*	Difference
Urban	87.3%	83.9%	-3.4%
Rural	12.7%	16.1%	3.4%

*Detail may not sum to total due to rounding.

To further assess the interaction between age group and strata, Table 2.7 shows the response rates by age group and strata. In general, as age increase, the overall response rates tend to increase as well. However, in the youngest age groups, 18-44, we see distinct differences in response rates between the enrollee strata. In particular, for those in MCO and FFS-PCCM we see lower response rates compared to those beneficiaries in the Full Duals and Disabled strata. The three older age groups show some differences in response rates, though not quite as divergent as the younger age groups. It should be noted that the Full Duals had the highest response rates across all five age groups.

[9] Hall, Susan A., Jay S. Kaufman, and Thomas C. Ricketts. "Defining urban and rural areas in US epidemiologic studies." *Journal of Urban Health* 83.2 (2006): 162-175.

Table 2.7: Distribution of Sample and Response Rates by Age Group and Enrollee Strata for the 2014-2015 NAM CAHPS Survey.

Age Group	Enrollee Strata Group	Sample Percent*	Response Rate
1: 18-44	Full Dual	4.4%	26.8%
	Disabled	9.6%	22.4%
	Managed Care	26.4%	11.9%
	FFS PCCM	19.9%	11.9%
2: 45-64	Full Dual	7.1%	37.5%
	Disabled	11.6%	35.6%
	Managed Care	4.9%	28.9%
	FFS PCCM	3.6%	28.8%
3: 65-74	Full Dual	5.3%	39.4%
	Disabled	0.2%	28.8%
	Managed Care	0.4%	33.2%
	FFS PCCM	0.9%	33.2%
4: 75+	Full Dual	4.7%	34.9%
	Disabled	0.0%	22.7%
	Managed Care	0.3%	29.3%
	FFS PCCM	0.7%	29.2%

*Detail may not sum to total due to rounding.

In conclusion from the tables presented above, the key findings are:

1. Full-Duals and Disabled beneficiaries were more likely to respond (Table 2.1), but these groups also have a higher portion of older adults who also were more likely to respond (Table 2.2).
2. Race and ethnicity did not appear to be a strong factor in response rate (Table 2.3), but this variable was based on what was captured from CMS data, and may not be an accurate indication of self-identified race and ethnicity.
3. Sex (Table 2.4), region (Table 2.5), and urbanicity (Table 2.6) did not appear to influence nonresponse.

3.0 Regression Analysis

To determine which characteristics are associated with the propensity to respond, a logistic regression analysis was completed to supplement the findings in Section 2 above. The focus of the logistic regression analysis is to determine which characteristics are most influential in whether a sampled Medicaid beneficiary responded to the NAM CAHPS survey.

Whether a sampled beneficiary responded was modeled with the following predictors: sex, race and ethnicity, urbanicity, stratum, age group, and census region, and with the following interaction terms: age by stratum, age by sex, age by race and ethnicity, age by region, age by

urbanicity, sex by race and ethnicity, and region by urbanicity. Note that as with Section 2 above, this analysis is limited to variables available on the sampling frame and for both respondents and nonrespondents. As such, some of these variables may not be as accurate as the same variables that are self-reported by respondents (e.g., race and ethnicity). A logistic regression model was fit using an iterative stepwise selection method that assessed inclusion and exclusion of variables in the model at each step.

Given the limited number of socio-demographic and geographic predictor variables available for the analysis, and the large number of cases used in this analysis, all single variables used in the model were significant, along with the stratum by age interaction term. Table 3.1 shows the results of the modeling, and the order in which each variable was included in the model. Of note, age group appears to be the strongest predictor, followed by stratum, in predicting likelihood of survey response.

Table 3.1: Results of Logistic Regression Analysis Modeling Propensity to Respond, in Order of Selection, for the 2014-2015 NAM CAHPS Survey.

Variable	Chi-Square	P-Value
Age Group	66,162.8	<.0001
Stratum	13,113.2	<.0001
Stratum*Age Group	3,150.0	<.0001
Urbanicity	2,563.1	<.0001
Sex	2,123.0	<.0001
Race and Ethnicity	801.5	<.0001
Census Region	749.0	<.0001

As seen from Table 3.1, age group has the most influence on the propensity to respond. For the purposes of the regression analysis, age group was divided into 4 groups: 18-44, 45-64, 65-74, and 75 or greater.

Across all four strata, 18-44 year olds had the lowest propensity to respond. Table 3.2 shows the odd ratios between the age groups by strata. Within each strata, the odds ratio was greater than 1 for every age group compared to the 18-44 age group, implying the older age groups had a higher propensity to respond in all four strata. These differences were most notable for the MCO and FFS-PCCM groups in which older age groups were more than three times more likely to respond compared to the youngest group. Across three of the four strata, the 65 – 74 year old group was the most likely to respond (i.e. highest odds ratio within each strata). The one exception was the Disabled stratum, where 45-64 year olds were most likely to respond.

As shown above in Section 2 and supported with this analysis, the older age groups were more likely to respond, and the strata associated with older adults (Full Duals, Disabled) were also more likely to respond. Table A.1 in the appendix shows the odds ratios between strata by age group, demonstrating that Full Duals were most likely to respond compared to other strata, within each of the age groups. The interaction term of age group and stratum was the third most

significant variable input in the model. It should be noted there was a statistically significant association between age group and stratum (Chi-Squared=557,778, p-value= <0.0001), indicating there is an interaction effect between the two variables.

Table 3.2: Odds Ratios by Age Group within Strata, for the 2014-2015 NAM CAHPS Survey.

Age Group within Strata	Odds Ratio	95% Confidence Interval
Full Duals		
18-44	1.00	Reference
45-64	1.61	1.57-1.65
65-74	1.75	1.70-1.79
75+	1.42	1.38-1.45
Disabled		
18-44	1.00	Reference
45-64	1.88	1.85-1.91
65-74	1.45	1.32-1.59
75+	1.08	0.87-1.35
Managed Care		
18-44	1.00	Reference
45-64	3.17	3.10-3.23
65-74	3.77	3.54-4.01
75+	3.20	2.96-3.45
FFS-PCCM		
18-44	1.00	Reference
45-64	3.14	3.06-3.21
65-74	3.89	3.72-4.06
75+	3.27	3.11-3.44

Table 3.3 below presents the odds ratios for the other variables included in the model (urbanicity, sex, race and ethnicity, and Census region). Those in urban areas were statistically less likely to respond than those in rural areas (odds ratio 0.75; 95% CI 0.74-0.76); and males were statistically less like to respond than females (odds ratio 0.75; 95% CI 0.74-0.76). Additionally, Hispanics (odds ratio 0.78; 95% CI 0.76-0.80) and Non-Hispanic Blacks (odds ratio 0.88; 95% CI 0.87-0.90) were less likely to respond than Non-Hispanic Whites. Those in the Midwest (odds ratio 1.22; 95% CI 1.20-1.24), South (odds ratio 1.13; 95% CI 1.11-1.14), and West (odds ratio 1.16; 95% CI 1.14-1.17) were more likely to respond than those living in the Northeast.

Table 3.3: Odds Ratios by Urbanicity, Sex, Race and Ethnicity, and Census Region, for the 2014-2015 NAM CAHPS Survey.

Variable	Odds ratio	95% Confidence Interval
Urbanicity		
Rural	1.00	Reference
Urban	0.75	0.74-0.76
Sex		
Female	1.00	Reference
Male	0.80	0.79-0.81
Race and ethnicity		
Non-Hispanic White	1.00	Reference
Non-Hispanic Black	0.88	0.87-0.90
Hispanic	0.78	0.76-0.80
Non-Hispanic Other	0.92	0.90-0.94
Missing	1.00	0.99-1.01
Census Region		
Northeast	1.00	Reference
Midwest	1.22	1.20-1.24
South	1.13	1.11-1.14
West	1.16	1.14-1.17

4.0 Benchmark Comparisons

Using the same socio-demographic variables described in prior sections, the distributions of survey respondents can be compared to available external benchmarks, primarily the CMS 2015 Statistics Reference Book[10], supplemented with estimates released from the Henry J. Kaiser Family Foundation (KFF). The CMS Statistics Reference Books are published every year based on current data about Medicaid and Medicare beneficiaries. However, the data published are for the broader Medicaid and Medicare population, including those outside the criteria established for the 2014-2015 NAM CAHPS survey. The KFF data allowed better comparisons for some of the characteristics of interest. In particular, the age group breakdown was more refined for 18-64 year olds than publicly available data from CMS[11]. The KFF data are widely reported, and use

[10] https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/CMS-Statistics-Reference-Booklet/Downloads/2015CMSStatistics.pdf

[11] KFF reports Medicaid enrollment by multiple age groups (0-18, 19-26, 27-44, 45-64, 65 years and over), while the CMS Statistics Reference Books only report age group by Under 21, 21-64, and 65 years and over. Since the NAM CAHPS survey

various data sources, including both CMS and state agency reporting. The CMS statistics were the default for comparison whenever possible, with the KFF data having more detailed information around age groups in particular.

The comparisons were run both unweighted and after applying final survey weights. As described in Section 1, the final sampling frame was representative of the targeted population but may not be representative of the adult Medicaid population as a whole. However, the benchmark comparison analysis described below was to analyze comparisons between the sampled beneficiaries and the general adult Medicaid population. Therefore, differences noted in this section may be due to differences between the target population and the general adult Medicaid population.

As a reminder, the goal of the NAM CAHPS survey was to have an equivalent number of respondents within each state and stratum combination. Thus, NORC drew the same number of sampled cases within each state and stratum, accounting for differential response rates between strata. As a result, states with larger adult Medicaid populations had a smaller sampling fraction than states with smaller adult Medicaid populations. For example, a larger proportion of eligible Medicaid beneficiaries were sampled from Rhode Island than from California. It is important to note this as the unweighted population distributions shown below will be reflective of the sample design implemented and not the targeted population.

The final weights for the 2014-2015 NAM CAHPS Survey accounted for the differential sampling fractions across strata and states. When the final weights are applied, most of the demographics align closer to the CMS benchmarks. While care was taken to make the sample as representative of the general adult Medicaid population as possible, there were limited variables deemed accurate for use during sampling. These limitations are noted with the results of each table where applicable.

The sample was selected by sorting by year of birth, along with sex and zip code, within each state and strata, therefore, allowing for a fairly representative sample as shown in Table 4.1a. The final column shows the difference between the weighted NAM CAHPS respondents compared to the KFF benchmark age group distributions. However, for the youngest age group, there were specific criteria applied to who was considered an eligible beneficiary for this survey. Specifically, beneficiaries with a family planning waiver, who were more likely to be in the 18-44 age group, were removed from the sampling frame, which impacted some of the age distributions.

Table 4.1a indicates that the youngest age group (19-26 years of age) had the largest differences compared to the benchmark distribution. These difference in the 19-26 age group distribution may have been impacted by the removal of those beneficiaries with a family planning waiver for the NAM CAHPS survey. As an additional step, further review was done for adults 27 years of age and older (i.e., all age groups KFF reports for with the exception of the youngest adult age group) to account for this sampling criteria that may have impacted the age group distribution comparisons. By limiting the data to those ages 27 years of age and older (see Table 4.1b), there

covered all adults 18 and older, the CMS age groups did not allow further understanding of response differences for those 21-64, while KFF allowed more refined comparisons.

is less discrepancy between the weighted respondents and KFF benchmark distribution. The difference between the two distributions shows that there are 3.5 percentage points more for those who are 45-64 years of age in the NAM CAHPS survey compared to the benchmark distribution. Meanwhile the oldest age groups show underrepresentation (-3.5 percentage points) compared to the KFF benchmark distribution even after accounting for weighting adjustments; though both of these differences would be considered minimal given the overall age group distributions. Given the differences noted between Table 4.1a and 4.1b distributions, this indicates that some, but not all, of the overall differences by age group are impacted by the eligibility criteria applied for the 2014-2015 NAM CAHPS survey. Note that these tables are limited to those ages 19 and above given the reporting groups from KFF.

Table 4.1a: Distribution of Respondents, Unweighted and Weighted, by Age Group versus CMS Benchmarks for the 2014-2015 NAM CAHPS Survey, Ages 19 and above.

Age Group	Respondents (Unweighted) *	Respondents (Weighted) *	Henry J. Kaiser Family Foundation[12]*	Difference (Weighted Respondents - Benchmark)
19-26	11.6%	15.8%	24.0%	-8.2%
27-44	24.38%	37.2%	34.2%	3.0%
45-64	37.75%	28.1%	23.2%	4.9%
65+	18.49%	16.9%	18.3%	-1.4%
Unknown	7.82%	2.0%	0.3%	1.7%

*Detail may not sum to total due to rounding.

Table 4.1b: Distribution of Respondents, Unweighted and Weighted, by Age Group versus CMS Benchmarks for the 2014-2015 NAM CAHPS Survey, Ages 27 and above.

Age Group	Respondents (Unweighted) *	Respondents (Weighted) *	Henry J. Kaiser Family Foundation[13]*	Difference (Weighted Respondents - Benchmark)
27-44	30.2%	45.2%	45.2%	0.0%
45-64	46.8%	34.2%	30.7%	3.5%
65+	22.9%	20.6%	24.1%	-3.5%

*Detail may not sum to total due to rounding.

In creating the sampling design for the survey, NORC assessed the existing race and ethnicity variables that could be leveraged for sampling. It was decided that those variables may not be

[12] KFF benchmarks retrieved from https://www.kff.org/medicaid/state-indicator/medicaid-enrollment-by-age/?currentTimeframe=0&sortModel=%7B%22colId%22:%22Location%22,%22sort%22:%22asc%22%7D.

[13] KFF benchmarks retrieved from https://www.kff.org/medicaid/state-indicator/medicaid-enrollment-by-age/?currentTimeframe=0&sortModel=%7B%22colId%22:%22Location%22,%22sort%22:%22asc%22%7D.

accurate, and would potentially vary in accuracy and completeness by state. Therefore, when selecting the sample, the sampling frame was sorted by zip code, with the assumption that race and ethnicity may be clustered geographically. Table 4.2 shows the distributions of race and ethnicity for the respondents, unweighted and weighted, compared to the benchmark. The race and ethnicity for the respondents was based on the self-reported race and ethnicity from the survey data, not from the sampling frame. While the sample was not directly selected by race and ethnicity, the weighted respondents' race and ethnicity distribution have minimal differences compared to the CMS benchmark distribution, where the final column of the table show differences ranging from -2.9 percentage points to 3.8 percentage points. There were more respondents who self-identified as Non-Hispanic White and fewer Hispanics compared to CMS benchmarks; however, these differences were mitigated with the use of the final survey weights. The final weights did result in a population with more non-Hispanic White and fewer non-Hispanic Black and Hispanic individuals.

Table 4.2: Distribution of Respondents, Unweighted and Weighted, by Race and ethnicity versus CMS Benchmarks for the 2014-2015 NAM CAHPS Survey.

Race and ethnicity	Respondents (Unweighted) *	Respondents (Weighted) *	2015 CMS Statistics Reference Book[14]*	Difference (Weighted Respondents - Benchmark)
1 - NH, White	56.7%	43.9%	40.1%	3.8%
2 - NH, Black	20.0%	19.2%	22.1%	-2.9%
3 - Hispanic	13.0%	24.1%	25.2%	-1.1%
4 - NH, Other	10.4%	12.8%	12.7%	0.1%

*Detail may not sum to total due to rounding.

The sampling frame was sorted by sex based on the sex available on the provided data, and as seen below in Table 4.3, aligns fairly well with the benchmark data. While males appeared to be less likely to respond to the survey compared to external benchmarks, the survey weights mitigated some of the differences. That being said, the final weighted respondents still produced results that underrepresented males compared to the CMS benchmarks (-4.5 percentage points).

Table 4.3: Distribution of Respondents, Unweighted and Weighted, by Sex versus CMS Benchmarks for the 2014-2015 NAM CAHPS Survey.

Sex	Respondents (Unweighted) *	Respondents (Weighted) *	2015 CMS Statistics	Difference (Weighted

[14] U.S. Department of Health and Human Services, United States, 2015 CMS Statistics Reference Booklet.

			Reference Book[15]*	Respondents - Benchmark)
Female	65.3%	62.9%	58.4%	4.5%
Male	32.9%	35.0%	41.4%	-6.4%
Unknown	1.7%	2.1%	0.2%	1.9%

*Detail may not sum to total due to rounding.

The respondents were categorized into four broad regions based on the nine CMS regions. It should be noted though that four states were not able to participate in this project (Alaska, New Hampshire, North Dakota, and Wisconsin). While the regional distributions were compared, the benchmark data includes states that were not part of the survey, which could affect the differences noted below. Unfortunately, state-level estimates for the 2015 CMS Statistics Book were not available, and as such the four states could not be removed from the benchmark totals. The final weighted data shows regional differences compared to the final CMS benchmarks (ranging from -6.6 percentage points to 6.4 percentage points), but it is difficult to determine if this is due to true differences or the exclusion of the four states noted above.

Table 4.4: Distribution of Respondents, Unweighted and Weighted, by Region versus CMS Benchmarks for the 2014-2015 NAM CAHPS Survey.

Region	Respondents (Unweighted)*	Respondents (Weighted) *	2015 CMS Statistics Reference Book[16]*	Difference (Weighted Respondents - Benchmark)
1 - Northeast	13.5%	21.5%	15.1%	6.4%
2 - Midwest	21.5%	16.8%	20.1%	-3.3%
3 - South	39.3%	33.4%	40.0%	-6.6%
4 - West	25.6%	28.3%	24.8%	3.5%

*Detail may not sum to total due to rounding.

In conclusion from the tables presented above, the key findings are:

1. Compared to benchmark distributions, the sampled respondents differed in the following ways
    a. Higher percent of 45-64 year olds;
    b. Higher percent of Non-Hispanic Whites and fewer Non-Hispanic Blacks and Hispanics;
    c. Higher percent of females;
    d. Higher proportion from the Northeast and lower proportion from the South.

[15] U.S. Department of Health and Human Services, United States, 2015 CMS Statistics Reference Booklet.

[16] U.S. Department of Health and Human Services, United States, 2015 CMS Statistics Reference Booklet.

2. Age group and region appeared to have more impactful discrepancies than race and ethnicity or sex.
    a. Of the demographics shown for comparison, the age group discrepancies would appear to have the most impact on being able to make generalizable conclusions from the 2014-2015 NAM CAHPS survey to the Medicaid population as a whole. NORC recommends limiting generalizations about the younger age groups, as these may not be as representative of all Medicaid beneficiaries.
    b. Sex also appears to have larger discrepancies compared to the benchmarks, with fewer males represented in the NAM CAHPS survey compared to the general Medicaid population.
    c. Non-Hispanic Blacks and Hispanics also appear underrepresented compared to the benchmarks.
    d. Care should be taken when comparing geographic distributions aggregating across states given not every state was included in NAM CAHPS survey.

5.0 Discussion

In general, the 2014-2015 NAM CAHPS survey respondents mimic the distribution from the sampling frame and known Medicaid beneficiaries with some key differences.

1. Younger age groups were less likely to respond than their older counterparts. Additionally, the criteria applied to be eligible for the survey limited the number of younger people that was included in the survey compared to the general Medicaid population. Therefore, care should be taken when making generalizations from the survey to both younger age groups and strata that are associated with younger age groups (i.e., MCO, FFS-PCCM).
2. Non-Hispanic Whites were more likely to respond to the survey and also made up a slightly larger proportion of respondents compared to the Medicaid benchmarks. While there are differences, the differences are minimal overall and likely do not impact the ability to generalize across these broad race and ethnicity categories.
3. Females were more likely to respond than males to the survey and were overrepresented in the NAM CAHPS survey compared to the external benchmarks. Depending on the nature of the analysis using the NAM CAHPS data, care should be taken when generalizing outcomes from the survey by sex.
4. Beneficiaries living in rural areas were more likely to respond compared to those in urban areas. There were no easily accessible external Rural/Urban benchmarks; therefore, it is difficult to determine how the weighted respondents Rural/Urban distribution compares to the general adult Medicaid population.
5. Care should be taken when generalizing for given sub-national geographic areas, as these may not be as representative of the Medicaid population in that area. This is at least partially due to four states which were unable to participate in the survey.
6. One limitation of the NAM CAHPS survey was the lack of detailed race and ethnicity information available on the sampling frame. Had this information been more robust, it could have been used for further sampling stratification, to ensure better race and ethnicity representation in the sample. While NORC was not able to explicitly sample by

race and ethnicity, Table 4.2 indicates there were no large deviations between the NAM CAHPS respondents and CMS benchmark distributions.

Authors: A. Elizabeth Ormson[17], Alison Laffan[17], Elsa Haile[18], Paul Guerino[19]

Suggested Citation: Ormson, A. E., Laffan, A., Guerino, P., and Haile, E., 2014-2015 Nationwide Adult Medicaid Consumer Assessment of Healthcare Providers and Systems (NAM CAHPS) Nonresponse Bias Analysis. Centers for Medicare & Medicaid Services. 2018.

[17] NORC at the University ofChicago.

[18] CMS Office of Minority

[19] Former employee, CMS Office of Minority

Appendix

Table A.1 shows the odds ratios for each strata within age groups, the converse of the odds ratios presented in Table 3.2. Across all age groups, the odds ratios were less than 1 for all strata compared to Full Duals. This implies the Full Dual stratum was more likely to respond to the NAM CAHPS survey than beneficiaries in other strata.

Table A.1: Odds Ratios by Strata within Age Group, for the 2014-2015 NAM CAHPS Survey.

Strata within Age Group	Odds Ratio	95% Confidence Interval
18-44		
Full-Duals	1.00	Reference
Disabled	0.80	0.78-0.82
Managed Care	0.36	0.35-0.36
FFS-PCCM	0.36	0.35-0.36
45-64		
Full-Duals	1.00	Reference
Disabled	0.93	0.92-0.95
Managed Care	0.70	0.68-0.72
FFS-PCCM	0.69	0.67-0.71
65-74		
Full-Duals	1.00	Reference
Disabled	0.66	0.60-0.73
Managed Care	0.77	0.72-0.82
FFS-PCCM	0.79	0.76-0.83
75+		
Full-Duals	1.00	Reference
Disabled	0.61	0.49-0.76
Managed Care	0.81	0.74-0.87
FFS-PCCM	0.82	0.78-0.86