The Star Treatment: Estimating the Impact of Star Ratings on Medicare Advantage Enrollments. Appendices

Michael Darden Department of Economics

Tulane University

Ian M. McCarthy Department of Economics

Emory University

May 2014

Appendix A: Star Rating Metrics

The star rating system can be divided into five domains, with the names of each domain, the underlying metrics in each domain, and the data sources for each metric changing over the years. The metrics and relevant domains for 2009 are listed in Table 1. In 2009 and 2010, a contract's overall star rating was based solely on Part C measures. Since 2011, the overall star rating now includes measures of Part D performance. Our focus on the 2009 and 2010 star rating program therefore avoids the 2011 and 2012 rating considerations specific to Part D plans. Our analysis also avoids additional computational difficulties introduced in the 2012 rating system that further obfuscate the overall star rating calculations (e.g., a weighted average of individual measures, with weights determined by the category of measure such as process measures with a weight of 1, access and patient experience measures with a weight of 1.5, or outcome measures with a weight of 3).

Table 1

Appendix B: Star Rating Calculations

Although the domains and individual metrics changed from year to year, the way in which overall star ratings were calculated was consistent in 2009 and 2010. The calculations follow in five steps, as described in more detail in the CMS technical notes of the 2009, 2010, and 2011 star rating calculations:

1. Raw summary scores for each individual metric are calculated as per the definition of the metric in question. As discussed in the text, these scores are derived from a variety of different datasets including HEDIS, CAHPS, HOS, and others. The resulting summary scores are observed in our dataset.

2. The summary scores in each metric are translated into a star rating. For most measures, the star rating is assigned based on percentile rank; however, CMS makes additional adjustments in cases where the distribution of scores are skewed high or low. Scores derived from CAHPS have a more complicated star calculation, based on the percentile ranking combined with whether or not the score is significantly different from the national average. The resulting stars for each individual metric are observed in our dataset.

3. The star values from each metric are averaged to generate an overall average rating for the contract, provided a minimum number of metric-level scores are available. For example, in 2009 and 2010, a ``Staying Healthy'' domain-level star was only calculated if the contract had a star value for at least 6 of the 12 individual measures.

4. Overall Part C summary scores are then calculated by adding to the overall average score an integration factor (i-Factor). The i-Factor is intended to reward consistency in a contract's quality across individual metrics, and is calculated as follows:

   (a) Derive the mean and variance of all individual metric summary scores for each contract.

   (b) Form the distribution of the mean and variance across contracts.

(c) Assign an i-Factor of 0.4 for low variance (below 30th percentile) and high mean (above 85th percentile), 0.3 for medium variance (30th to 70th percentile) and high mean, 0.2 for low variance and relatively high mean (65th to 85th percentile), and 0.1 for medium variance and relatively high mean. All other contracts are assigned an i-Factor of 0.

5. Overall Part C star ratings are then calculated by rounding the overall summary score to the nearest half-star value.

We do not observe the i-Factors in the data. We therefore replicated the CMS methodology, ultimately matching the overall star ratings for 98.8% and 98.5% of the plans in 2009 and 2010, respectively. As discussed in the text, plans for which we were unable to replicate star ratings were dropped from the analysis. Note also that star ratings are based on data from at least the previous calendar year and sometimes further back depending on ease of access from CMS. New plans therefore do not have a star rating available, nor was a star rating for such plans provided to beneficiaries.

Tables 2 and 3 presents example calculations of the overall summary score and resulting star values for 5 contracts in 2009. The table lists the summary scores for the individual metrics along with the corresponding star values, each of which are observed in the raw data. The high mean and low mean thresholds for i-Factor calculations were calculated to be 3.6667 and 3.2381, respectively. Similarly, the high variance and low variance thresholds were 1.3462 and 1.0362, respectively.

Table 2 and 3

The calculations for each contract in Table 2 are discussed individually below:

1. Contract H0150: With a mean star value of 2.583 and a variance of 0.879, the contract received an i-Factor of 0 (due to a low mean), which provided an overall summary score of 2.583 and a star rating of 2.5.

2. Contract H0151: With a mean star value of 2.667 and a variance of 0.8, the contract received an i-Factor of 0 (again from a low mean), which provided an overall summary score of 2.667 and a star rating of 2.5, just 0.083 points away from receiving a 3-star rating.

3. Contract H1558: With a mean star value of 3.967 and a variance of 1.275, the contract received an i-Factor of 0.3 (high mean and medium variance), which provided an overall summary score of 4.267, just 0.0167 above the 4.25 threshold required to round up to a 4.5-star rating.

4. Contract H0755: With a mean star value of 3.5278 and a variance of 1.285, the contract received an i-Factor of 0.1 (relatively high mean and medium variance), which provided an overall summary score of 3.6278 and a star rating of 3.5.

5. Contract H1230: With a mean star value of 3.694 and a variance of 1.018, the contract received an i-Factor of 0.4 (high mean and low variance), which provided an overall summary score of 4.094 and a star rating of 4.0.

Appendix C: Data

Our analysis merges publicly available data from several sources. We start with enrollment and plan characteristic data at the county/month/year/contract/plan level from June of 2008 through December of 2011.[1] Each contract offers a menu of plans, and ours is the first analysis of which we are aware to model enrollment trends at the plan level. For contracts that only offer one plan, we code the plan indicator as a 1.

As our starting point, we construct a census of all counties in the United States from 2008 to 2011 and all MA contracts that were approved to operate within each county, month, and year.[2] Following CMS and previous work, we define the county as the relevant market area, and the service area files identify the MA choice set. Between June 2008 and December 2011, the Service Area files have 15,834,943 county/contract/month/year observations. We exclude 252,186 observations from U.S. territories. 6,496 of the remaining observations are duplicates at the county/contract/month/year level. 83.96% of these duplicates are in Los Angeles county, where the Service Area files, enrollment data, and contract information are further split by Social Security Administration (SSA) code. The duplicates are split evenly across years and months in our sample. Because the same contracts are available in each SSA within LA county, and because CMS considers the county to the be relevant market area, and because we do not have data at SSA level, we drop all but one record of these SSAs. In all other data sources, we aggregate data from the SSA to the county level. After these adjustments, we have 15,576,261 county/contract/month/year observations.

Contract information and contract/plan enrollment data are in separate CMS data files; thus, we merge these together first prior to merging this information into the Service Area files. Each file lists all contract/plans in all counties that have positive enrollment for the month and year of that

[1] Enrollment and plan data available at: http://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/MCRAdvPartDEnrolData/Monthly-Enrollment-by-Contract-Plan-State-County.html. Data are not available for July of 2008.

[2] Service Area Data available at: http://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/MCRAdvPartDEnrolData/MA-Contract-Service-Area-by-State-County.html.

file. However, enrollment is left missing if the contract/plan has fewer than 11 enrollees. Rather than impute the enrollment for these rows, we drop contract/plans with enrollment less than 11. Between June 2008 and December 2011, there are 90,032,404 contract/plan/county/month/year observations; after dropping rows with missing enrollment, we are left with 5,040,619 contract/plan/county/month/year observations. Of these, we drop 2,770 observations that are missing a FIPS code and drop all observations from U.S. Territories. We are left with 4,932,975 observations, of which 1,397 are duplicates - again mostly due to the Los Angeles County SSA split. We aggregate enrollment counts for these to the county level and drop the duplicates. Thus, the final count is 4,931,578 contract/plan/county/month/year observations.

The contract/plan specific information (other than quality information) are in separate files, and data are available for all years in the enrollment and service area files.[3] For the period between June 2008 and December 2011, these files contain 250,923 unique contract/plan/month/year observations.

To merge these three files together, we first merge the contract information to the enrollment files. The merge is on the contract/plan/month/year dimension because the contract characteristics are not specific to counties. Of the 250,923, contract/plan/month/year observations in the contract data, 65,875 do not have a corresponding record in the enrollment data due to either: a) enrollment$<$11; or b) a missing FIPS code. We drop these contract/plan observations leaving 4,931,578 contract/plan/county/month/year observations. Importantly, there are no observations in the merge that are in the enrollment data that do not exist in the contract data. This augmented enrollment dataset is then merged to the Service Area file. Note that we must use the Service Area File because the enrollment records contain contract/plans that are not approved to operate in their respective county, and thus not part of a prospective enrollee's choice set. The merge with the service area file is on the contract/county/month/year dimension. There are many more approved

[3]Contract/plan information available at: http://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/MCRAdvPartDEnrolData/Monthly-Enrollment-by-Contract-Plan-State-County.html.

contracts than appear in the enrollment data.

Of the 4,931,578 contract/plan/county/month/year enrollment observations, only 1,557,531 have a corresponding entry in the service area file. However, 99.21% of those that do not have a corresponding record in the service area files are Medicare Prescription Drug or Employer PDP plans. The remaining unmatched observations (24,158 or 1.69%) can only be those beneficiaries in a contract/plan with at least 10 enrollees that kept their plan despite moving counties. We drop these PDP contract/plans, as well as any special needs and employer specific plans.[4] We drop all approved contract/plans with no enrollees (or none above 10 enrollees). The final contract/plan/county/month/year dataset contains 1,403,355 observations. From here on, we call these data the plan data.

Census data are from the American Community Survey for 2006-2010 in wide format. Merging these to the plan data on the FIPS code, all observations in the plan data have a corresponding county entry in the census data, but we drop the 179 counties (0.01%) that have no contract/plans with significant ($>$10) enrollment.

For each year from 2009-2011, CMS provides information at the contract level on a.) the overall summary star, b.) the star level for each domain (e.g., staying healthy), c.) individual measure stars (e.g., breast cancer screening), d.) the continuous individual measure value, and e.) the cut points for each individual star threshold.[5] We construct a Stata dataset for each of a.-d. for each year and merge together on the contract identifier. All merges are perfect such that we have yearly quality datasets. There is information on 754, 709, and 575 contracts in 2009, 2010, and 2011, respectively. Data for 2008 have all of the components above, but lack the overall summary measure for each contract. We proceed similarly for 2008, with 413 contracts having information.

[4]Our final sample consists only of HMOs, PPOs, and private fee-for-service (PFFS) contracts. We do not adjust the count of county Medicare eligibles for special needs or employer specific plan enrollment when we construct the share of individuals in the outside option $j = 0$ because nothing prevents these enrollees from switching to a traditional MA plan. Nonetheless, our results are unchanged when excluding these enrollees from the relative market share calculations.

[5]Contract-level quality data available at: http://www.cms.gov/Medicare/Prescription-Drug-Coverage/PrescriptionDrugCovGenIn/PerformanceData.html.

Appending these data together yields a dataset with 848 unique contracts and 2451 contract/year observations.

We merge these to the plan data and, of the 1,403,355 plan data observations, 1,353,559 have a contract/year record in the quality data. There are 452 contract/year observations in the quality data with no corresponding enrollment record because of enrollment <11, and these are dropped. We keep all enrollment records and create a dummy variable that equals one if there is no quality information.

Next, we merge the 1,403,355 plan data observations on contract/plan/county/year to premium information.[6] For 2008-2011, we have information for 650 unique counties and 331,354 observations at the contract/plan/county/year level after dropping 359 observations that were labeled as statewide. We drop 346 of these observations that are duplicates at the contract/plan/county/year level. Unfortunately, the premium data do not contain the state/county FIPS code, only the name of the state and county. We merge on state/county names to a file from the Census Bureau that maps state county names from 2011 to 2011 FIPS codes.[7] We are able to match all of the 330,959 contract/plan/county/year observations to a state/county FIPS code (there are 5 FIPS codes in the FIPS code registry that do not correspond to a state/county row in the premium data). Of these records, 324,952 observations have a nonmissing value for the plan premium. Of the 1,403,355 contract/plan/county/year observations, 1,024,074 have a corresponding row in the premium data. For those without a premium data row, we construct a dummy variable that indicates the missing value.

The final file from which we merge information is the Medicare county level penetration files.[8] This provides information on the number of individuals that are eligible in a county, and the num-

[6]Data on plan premiums available at: http://www.cms.gov/Medicare/Prescription-Drug-Coverage/PrescriptionDrugCovGenIn/index.html?redirect=/PrescriptionDrugCovGenIn/.

[7]County names and FIPS codes available at: http://www.census.gov/popest/about/geo/codes.html.

[8]MA penetration data available at:http://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/MCRAdvPartDEnrolData/MA-State-County-Penetration.html.

ber that are enrolled in an MA plan. There are 128,907 records at the county/month/year level. 1,403,351 of the contract/plan/month/year observations merge perfectly (only one contract in 2008 has no record in the penetration file). 12,047 of the 128,907 records in the penetration file do not have a match in the plan data. The discrepancy is that, while a county may have more than 10 enrollees, and thus a record in the penetration files, a contract/plan may not, and is thus omitted.

The final sample for 2008-2011, at the contract/plan/county/month/year level includes 1,403,351 observations. When considering just 2009, we have records for 466,239 contracts/plan/county/months. We then restrict our attention to HMOs, PPOs, and PFFS contract/plans, and we average over a year, which yields 41,730 contract/plan/counties. Our county level information is taken from October of the prior year. For 2010, we have 38,414 contract/plan/counties. Finally, we drop all observations with missing quality or premium data. Our estimation samples are 20,437 and 17,130 for 2009 and 2010, respectively.

Appendix D: Tables

Table 1: **Domains, Metrics, and Data Sources for 2009 MA Star Rating Program**[a]

Staying Healthy	**Getting Timely Care from Doctors**	**Plan Responsiveness and Care**	**Managing Chronic Conditions**	**Handling of Appeals**
Breast Cancer Screening (HEDIS)	Access to Primary Care Doctor Visits (HEDIS)	Getting Appointments and Care Quickly (CAHPS)	Osteoporosis Management (HEDIS)	Plan Makes Timely Decisions about Appeals (IRE)
Colorectal Cancer Screening (HEDIS)	Follow-up Visit within 30 Days of Discharge after Hospital Stay for Mental Illness (HEDIS)	Overall Rating of Health Care Quality (CAHPS)	Diabetes Care - Eye Exam (HEDIS)	Reviewing Appeals Decisions (IRE)
Cardiovascular Care - Cholesterol Screening (HEDIS)	Doctor Follow-up for Depression (HEDIS)	Overall Rating of Health Plan (CAHPS)	Diabetes Care - Kidney Disease Monitoring (HEDIS)	
Diabetes Care - Cholesterol Screening (HEDIS)	Getting Needed Care without Delays (CAHPS)	Call Answer Timeliness (HEDIS)	Diabetes Care - Blood Sugar Controlled (HEDIS)	
Glaucoma Testing (HEDIS)		Doctors Who Communicate Well (CAHPS)	Diabetes Care - Cholesterol Controlled (HEDIS)	
Appropriate Monitoring of Patients Taking Long-Term Medications (HEDIS)		Customer Service (CAHPS)	Antidepressant Medication Management (HEDIS)	
Annual Flu Vaccine (CAHPS)			Controlling Blood Pressure (HEDIS)	
Pneumonia Vaccine (CAHPS)			Rheumatoid Arthritis Management (HEDIS)	
Improving or Maintaining Physical Health (HOS)			Testing to Confirm COPD (HEDIS)	
Improving or Maintaining Mental Health (HOS)			Continuous Beta Blocker Treatment (HEDIS)	
Osteoporosis Testing (HOS)			Improving Bladder Control (HOS)	
Monitoring Physical Activity (HOS)			Reducing the Risk of Falling (HOS)	

[a]Description of domains and additional details available at www.cms.gov/Medicare/Medicare.html. Data source for CMS calculations provided in parenthesis.

Table 2: **Star Rating Calculation Examples**

	Stars					Raw Scores				
	H0150	H0151	H1558	H0755	H1230	H0150	H0151	H1558	H0755	H1230
Breast Cancer Screening	2	2	5	4	5	59	57	87	75	87
Colorectal Cancer Screening	2	3	4	5	4	35	45	62	71	59
Cardiovascular Care - Cholesterol Screening	3	3	4	4	5	79	81	93	90	96
Diabetes Care - Cholesterol Screening	3	2	4	4	4	77	74	88	92	94
Glaucoma Testing	3	3	5	5	4	60	60	84	76	73
Appropriate Monitoring for Long-term Medications	4	3	5	4	2	90	88	93	90	82
Annual Flu Vaccine	3	2	5	4	5	67	66	87	77	84
Pneumonia Vaccine	3	2	5	3	4	67	63	80	68	77
Improving or Maintaining Physical Health	3	3	3	3	3	60	54	59	60	55
Improving or Maintaining Mental Health	3	3	3	3	3	81	78	81	82	80
Osteoporosis Testing	1	2	3	3	3	56	58	68	68	71
Monitoring Physical Activity	3	3	3	3	3	46	41	44	48	51
Access to Primary Care Doctor Visits	4	4	5	5	4	94	92	99	97	89
Follow-up after Hospital Visit for Mental Illness	3	2		4	5	46	41		72	77
Doctor Follow-up for Depression	1	1		1	2	5	3		20	22
Getting Needed Care without Delays	3	5	5	3	3	83	88	90	86	83
Getting Appointments and Care Quickly	1	2	5	4	3	68	72	83	77	75
Overall Rating of Health Care Quality	3	3	5	4	3	84	85	90	86	85
Overall Rating of Health Plan	4	4	5	3	4	86	87	92	86	87
Call Answer Timeliness	4	2	4	3	5	83	72	84	81	96
Doctors Who Communicate Well	3	4	5	4	4	90	91	93	91	91
Customer Service	3	3	5	3	3	88	87	92	88	86
Osteoporosis Management	1	1		1	2	17	16		19	28
Diabetes Care - Eye Exam	3	3	5	5	5	53	55	82	79	91
Diabetes Care - Kidney Disease Monitoring	3	3	3	4	5	76	77	77	85	97
Diabetes Care - Blood Sugar Controlled	2	2	4	5	4	53	55	82	87	83
Diabetes Care - Cholesterol Controlled	2	2	4	5	4	33	30	58	63	59
Antidepressant Medication Management	2	2		2	5	44	40		43	63
Controlling Blood Pressure	1	2	4	5	4	33	51	63	68	62
Rheumatoid Arthritis Management	2	3		3	3	68	71		73	75
Testing to Confirm COPD	2	2	2	2	2	24	21	33	32	30
Continuous Beta Blocker Treatment	3	2		2	4	79	69		73	85
Improving Bladder Control	2	2	2	2	2	37	34	39	38	37
Reducing the Risk of Falling	3	3	3	5	4	55	55	55	63	61
Plan Makes Timely Decisions about Appeals	4	4	1	4	5	86	88	43	91	100
Reviewing Appeals Decisions	1	4	3	3	3	66	86	79	77	77

Table 3: **Star Rating Calculation Examples, Cont.**

	H0150	H0151	H1558	H0755	H1230
Mean Summary Score	2.5833	2.6667	3.9667	3.5278	3.6944
Variance Summary Score	0.8786	0.80	1.2747	1.2849	1.0183
i-Factor	0	0	0.3	0.1	0.4
Summary Score	2.5833	2.6667	4.2667	3.6278	4.0944
Star Rating	2.5	2.5	4.5	3.5	4