Title
Evaluating the relationship between health plan characteristics and CMS quality scores

Permalink
https://escholarship.org/uc/item/58f7v8c9

Author
Suh, Joon Won

Publication Date
2021

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Evaluating the relationship between health plan characteristics and

CMS quality scores

A thesis submitted in partial satisfaction of the

requirements for the degree Master of Applied Statistics

by

Joon Won Suh

2021

ABSTRACT OF THE THESIS

Evaluating the relationship between health plan characteristics and

CMS quality scores

by

Joon Won Suh

Master of Applied Statistics

University of California, Los Angeles, 2021

Professor Hongquan Xu, Chair

The thesis uses multivariable linear regression to determine how different factors related to a

health plan impact its ability to provide its members with high quality medical care and customer

service as measured by the CMS, the federal government agency which administers the

Medicare program. The thesis finds that past CMS scores positively affect current CMS scores,

indicating that organizations can build competencies that improve CMS scores that carry

forward to future periods. In addition, it finds a negative relationship between the size of the

managed care market and CMS scores. The study finds a positive relationship between a plan's

level of focus on the Medicare Advantage business (percent of its members in the Medicare

Advantage plans) and CMS scores. Finally, the study presents a case for expanding the usage

of statistics and advanced analytics to help health plans evaluate their strategic options and

make more well-informed strategic and operational decisions.

The thesis of Joon Won Suh is approved.

Chad Hazlett

Mahtash Esfandiari

Hongquan Xu, Committee Chair

University of California, Los Angeles

2021

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

GLOSSARY

Term/Acronym	Definition
Center for Medicare and Medicaid Services (CMS)	Federal agency within the U.S. Department of Health and Human Services (HHS) that administers the Medicare program and other programs.
CMS Star Ratings	Annual incentive program created by CMS that awards health plans ratings from 1 to 5 stars based on their quality performance on its designated metrics. A 4 or 5-star designation is highly coveted by health plans because it can be used for new member market, and because those who receive this designation is awarded a significant bonus payment for the year from CMS.
Fee-for-service	Traditional method of payment where providers are paid for their services. This method is very costly, and CMS is actively working to transition to Medicare Advantage and other reimbursement programs to control spiraling medical costs.
Fully funded	Risk-based medical insurance plans. The health plan assumes financial risk for the member's medical care and contracts with hospitals and physician providers to share risk. The Medicare Advantage program is an example of a fully funded plan.
Health plan	Company that underwrites, sells, and administers the health insurance for individuals and businesses. Health plans also contract with government agencies including CMS to provide insurance for their beneficiaries.
Managed care	For this thesis, managed care and fully funded is defined in the same way and will be used interchangeably.
Medicare	National health insurance program in the U.S., administered by CMS. It primarily provides health insurance for Americans aged 65 and over.
Medicare Advantage (MA)	A type of health insurance product that provides Medicare benefits through a private-sector health insurer. The insurer is paid a fixed monthly rate by the government, and contracts with providers to share financial risk for the care of their assigned members.
Providers	In healthcare, providers are large systems of hospitals, other healthcare delivery sites, physicians and other clinicians who deliver healthcare to patients. Local examples of providers include Cedars-Sinai, a medical center and UCLA Medical Group, a large medical practice with thousands of physicians and other clinicians.
Special Needs Plan (SNP)	A specialized version of the Medicare Advantage (MA) product that also provides additional enhanced medical services to seniors that with greater healthcare needs because they are frail or have serious health conditions.

CHAPTER 1
Introduction

The immediate goal of this thesis to use multivariable regression to identify possible relationships between business strategies and their ability to help health plans improve the quality of care and services that they provide for their Medicare members. The broader goal is to provide support for the argument that health plans can use statistics and machine learning to improve their business strategies and optimize the returns on their investment by using a more quantitative approach.

The thesis uses healthcare data that is compiled and shared by the U.S. Government's Center for Medicare and Medicaid (CMS) Program, which administers the Medicare insurance program. Specifically, the data used is for year 2020 from the CMS Stars Program, which tracks 33 healthcare measures and is reported from each participating health plan for the Medicare Advantage members.

1.1 Medicare Program

In the United States, the federal government administrates the Medicare program, a health insurance program for over 60 million people [1]. The Medicare program is managed by the Center for Medicare and Medicaid Services (CMS). In year 2019, the U.S. government spent $787 billion on Medicare benefits [1], making Medicare the second largest federal program behind Social Security ($1 trillion).

Given the enormous cost and social impact of the Medicare program, the U.S. government has implemented strategies to control costs and improve quality. One strategy is to

spread adoption of its Medicare Advantage (MA) program, a managed care program focused on controlling medical costs and improving quality of care. As of 2019, the MA program served 20 million people and accounted for 34% of costs for the Medicare program.

1.2 Medicare Advantage (MA) Program

CMS offers the Medicare Advantage (MA) program for its beneficiaries (primarily seniors). Medicare Advantage is a managed care health insurance product, sold to beneficiaries and administered by commercial health plans. One of the key features of the MA product is that healthcare providers (e.g., hospitals and physician groups) accept reimbursement terms where they accept financial risk and are paid based on the quality of care they provide.

1.3 CMS Stars Program

For the Medicare Advantage program, CMS requires participating health insurance plans to participate in its CMS Stars program. In the CMS Stars program, participating health insurance plans collect health data on their members on 33 key metrics defined by CMS. Once this is reported to CMS by each plan, CMS analyzes the data, comparing each plan's performance on each metric against those of all other health plans.

CMS provides each plan with an overall star rating between 1 star (worst rating) to 5 stars (best rating). Every year, plans which score 4 or more stars receive a bonus payment from CMS. The potential bonus payment is roughly 5% of the health plan's payments from CMS. Given that for many health plans, net earnings represent only 2% of their revenues (e.g., Centene, 2019), the CMS bonus is a significant sum with material financial consequences.

CMS Stars program has a separate set of measures for its medical plan (aka Part C) plan and for its prescription benefit plan (aka Part D). This thesis only evaluates the CMS scores for its medical plan (Part C).

1.4 Health Plan Characteristics and CMS Stars Program Performance

The thesis uses multivariable linear regression to determine how different factors impact CMS Star Scores on its medical plan product (Part C). For this study, we do not consider the measures for its prescription benefit (Part D) program. The thesis studies the following health plan attributes or strategies shown in Table 1 below. The goal is to determine what, if any, associations exist between any of these topic areas and its CMS Star Scores.

Table 1: Topic Areas and Factors Studied by Linear Regression

Topic Area	Factor Studied
Organization's attributes	• For-profit vs non-profit status of health plan • Private or publicly traded health plan • Provides Special Needs Plan (SNP) • Organizational type • Operating in an area with a major disaster
Industry structure: Market penetration of managed care	• Market size • Market penetration of managed care
Company strategy: M&A	• Health plan size/scale • Health plan market share
Company strategy: Business focus	• Health plan focus on the Medicare Advantage business • Health plan focus on the broader managed care business
Company strategy: Build core capabilities	• CMS 5 Star score from previous years

CHAPTER 2
Datasets

2.1 CMS Dataset

CMS quality score data for 2020 is downloaded from the CMS website [3]. The data is

organized with each row being defined by a Medicare Contract ID. The Medicare Contract ID

uniquely identifies a health plan's Medicare Advantage insurance product that is sold in a

specific geographical market. The figure below provides a partial snapshot of the data. On the

first data row, Contract ID value of H2773 is for a Medicare Advantage product provided by a

company called Quality Health Plans for markets in New York state.

Figure 1: Screenshot Showing Part of Medicare CMS Stars Data File

	A	B	C	D	E	F	I	J	K	L
4	Contract ID	MCO	Org_Type	profit_status	private_public	Org_Name	SNP	DisasterPerc2017	DisasterPerc2018	Enough_Data_C
5	H2773	800	Local CCP	For-profit	Private	QUALITY HEALTH PLANS	1	0	0	Y
6	H0351	1543	Local CCP	For-profit	Public	Allwell	1	0	0	Y
7	H1436	1543	Local CCP	For-profit	Public	Allwell	1	0	16	Y
8	H3071	1906	Local CCP	For-profit	Private	Community Care Alliance of Illinois	0	0	0	Y
9	H4853	1894	Local CCP	For-profit	Private	Bright Health	0	0	0	Y
10	H7326	206	Local CCP	For-profit	Public	WellCare	0	0	0	Y
11	H7680	492	Local CCP	For-profit	Private	Prominence Health Plan	0	0	0	Y
12	H8908	1109	Local CCP	For-profit	Public	Humana	1	0	0	Y
13	H9276	1543	Local CCP	For-profit	Public	Allwell	0	100	0	Y

There are 743 records (unique Contract IDs) in the 2020 file provided by CMS. However,

CMS does not provide a Part C Star score for contracts in the following cases: (a) plan offers

Part D but does not offer a Part C product, (b) the plan is too new and is exempted from

providing data, or (c) the plan is exempted because it has too few members to provide sufficient

data. After these records are excluded in this study, the data used for the analysis has 409

records.

In addition to providing CMS Stars scores, CMS data also provides additional information on each health plan for evaluation during the statistical modeling such as by organization type, for-profit status, whether it also offers a special needs plan, and whether or not it operates in an area that experienced major disasters in previous years.

The dependent variable used in this study is an overall CMS score which is a weighted average of the 33 individual CMS metrics. A histogram of the overall CMS score is shown below. The 33 metrics, definitions and weights used to calculate this are in Appendix A. To compute an overall average score, CMS assigns different weights to the 33 individual CMS metrics to reflect the relative importance of each metric. The metric covering health plan quality improvement activities has the highest weight (5), followed by metrics that improve health outcomes (3), and followed by metrics that capture patients' experiences and complaints and overall access to services (1.5). Individual process metrics such as the metric that captures overall participation of members in the annual flu vaccine have the lowest weight (1). Appendix A has the list of metrics and the five categories mentioned above.

Table 2: Descriptive Statistics of Calculated 2020 Overall CMS Scores (Dependent Variable)

Descriptive Statistic	Value
Count	409
Mean	3.63
SD	0.44
Min	2.29
Max	4.66

Table 3: Values and Frequency of Occurrence for Key Factors

For-Profit Status		For Special Needs Plan		For Private/Public Organizations	
Value	Frequency	Value	Frequency	Value	Frequency
For-profit	265	SNP (0 = No)	219	Government	4
Not-for-profit	144	SNP (1 = Yes)	190	Mutual company	10
Grand Total	409	Grand Total	409	Private	196
				Public	199
				Grand Total	409

Figure 2: Histogram of Calculated 2020 Overall CMS Score (Dependent Variable)
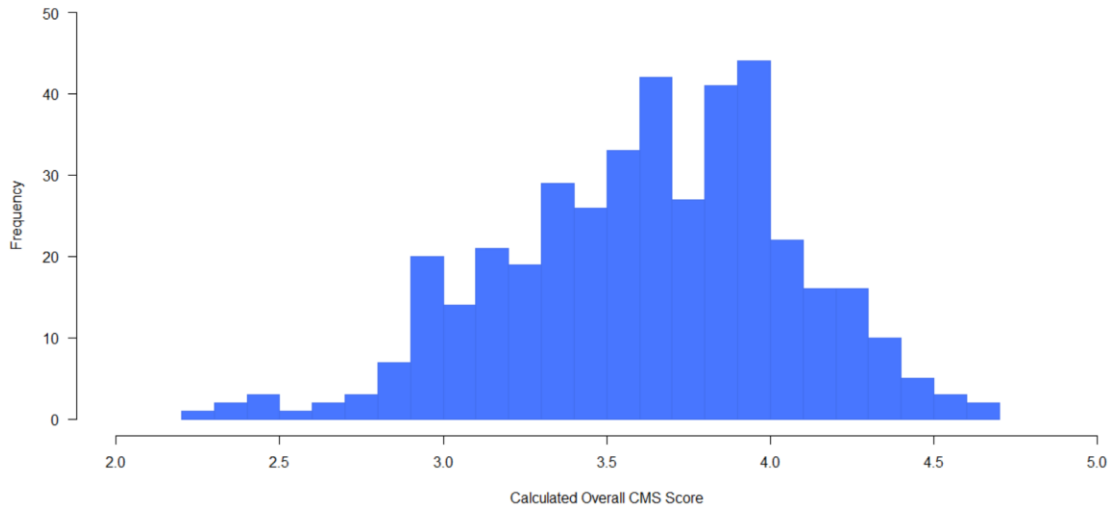
Figure 3: Normal Q-Q Plot of Calculated 2020 Overall CMS Score (Dependent Variable)
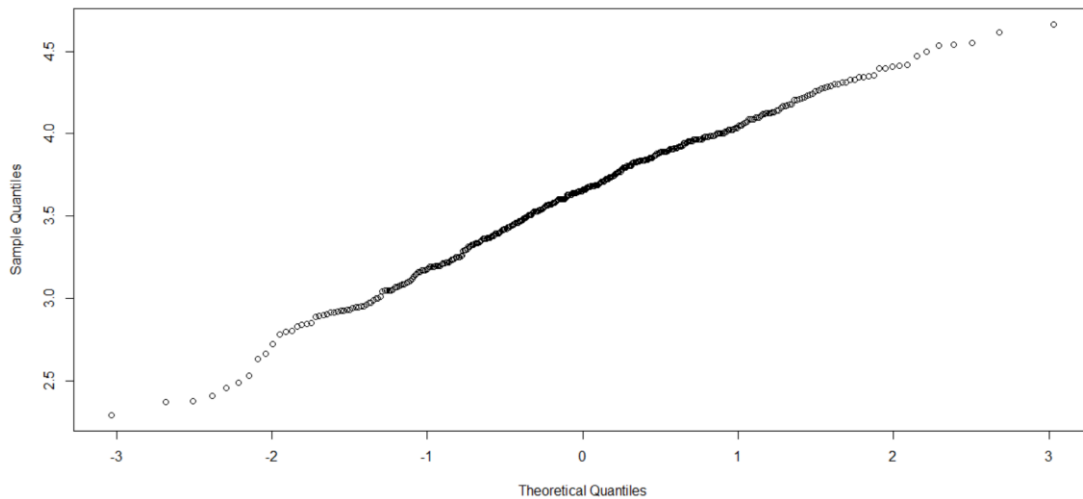
Figure 3 shows that the normality appears to be reasonable, with the exception of a few low outliers. The Anderson-Darling test is conducted on the outcome metric, creating a p-value of 0.009395 (under 0.05). The p-value, which would indicate that it is not a normal distribution, could be misleading due to the large sample size or the existence of outliers. Outliers do not represent a high percentage of the data. Removal of outliers in subsequent steps does not change the results of the model.

Figure 4: Symbox Transformation of Calculated 2020 Overall CMS Score (Dependent Variable)
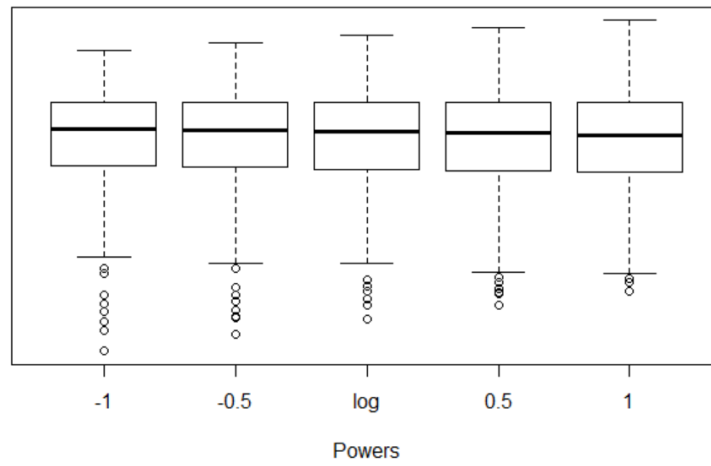
R symbox function is used to generate box plots with different power transformations of the dependent variable. Based on an assessment, the rightmost transformation (power is 1) is determined to have the most symmetric distribution and is used in the modeling. If the result shows that power 1 is the best transformation, then no transformation is needed.

2.2 Membership Data

Membership data is sourced from a third-party industry data source and is used to provide additional information on the health plan's size as well as market share information. Below is a data dictionary for the data used for the linear regression analysis.

Figure 5: Data Dictionary

Variable Name	Descriptive Name	Description	Source
Organization's Attributes			
contract_id	Medicare Contract ID	Value which uniquely identifies a MA product that is sold by a health plan in a define market	CMS
org_type	Organization Type	Additional information on the health plan. "1976 Cost" – Legacy MA health plan that is being phased out by CMS because health plans do not assume financial risk "Local CCP" – Local Coordinated Care Plan "MSA" – Medical Savings Account "PFFS" – Plan provided directly by employer or union for their employees/retirees "Regional CCP" – Local Coordinated Care Plan	CMS

profit_status	Profit Status	"For-profit" – health plan is a for-profit entity "Not-for-profit" – health plan is a not-for-profit entity	Online research
Private_public	Private Public	"Private" – health plan is not publicly traded "Public" – health plan is publicly traded	Online research
DisasterPerc2017	Disaster Percentage 2017	Percentage of MA enrollees for the plan who resided in a disaster area defined by CMS in 2017	CMS
DisasterPerc2018	Disaster Percentage 2018	Percentage of MA enrollees for the plan who resided in a disaster area defined by CMS in 2018	CMS
SNP	SNP (Special Needs Plan)	Flag which indicates whether or not the contract/MA product is a Special Needs Plan product which offered enhanced medical care to a sicker, more frail population: 1 = Yes 0 = No	CMS
Stars_2020_Calc_C	Calculated CMS Stars Score for 2020	Weighted average of 33 individual quality metric scores. Individual metrics and their weights are provided by CMS and shown in Appendix A.	Calculated based on CMS data
Contract_MA	Contract MA Members	Total number of members in Medicare MA program in the Medicare contract (i.e., Contract ID)	Membership data
Industry Structure			
Mkt_Tot	Market Total Members	For each health plan, provides the total number of members for all health plans in all of the states where the health plan does business	Membership data
Mkt_FF	Market Fully Funded Members	For each health plan, provides the total number of fully funded members for all health plans in all of the states where the health plan does business	Membership data
Mkt_MA	Market Medicare Advantage Members	For each health plan, provides the total number of Medicare Advantage members for all health plans in all of the states where the health plan does business	Membership data
Mkt_Tot_XMA	Market Total Members, Excluding Medicare Advantage Members	Market Total Members – Market Medicare Advantage Members	Derived
Mkt_FF_XMA	Market Fully Funded Members, Excluding Medicare Advantage Members	Market Fully Funded Members – Market Medicare Advantage Members	Derived
Mkt_FF_Bin	Market Fully Funded Members, Binned	Binned categories based on Mkt_FF: • "<10M" – under 10 million members • "10M – 100M" – greater than 10 million to 100 million members • ">100M" – greater than 100 million members	Derived
FF_Perc_Mkt	Fully Funded as a Percent of Market Size	For each health plan and the markets it serves, provides the percentage of fully funded members in the market as a proportion of the total number of members in the market: Mkt_FF / Mkt_Total	Derived
Factor Type: Health Plan M&A			

All_Total	All Members	Total number of members across all insurance products for the health plan associated with the Medicare contract	Membership data
All_Total_Bin	All Total Members Binned	Binned categories based on All_Total: • "<100k" – under 100,000 members • "100k – 1M" – from 100,000 to 1 million members • "1M – 10M" – greater than 1 million to 10 million members • ">10M" – greater than 10 million members	Derived
All_FullyFunded	All Fully Funded Members	Total number of members across all insurance products where the health plan associated with the Medicare contract assumes financial risk	Membership data
All_MA	All MA Members	Total number of Medicare Advantage members for the health plan. This is a subset of members in All_FullyFunded.	Membership data
All_Total_XMA	All Members Excluding MA Members	Total number of members across all insurance products excluding MA members for the health plan associated with the Medicare contract: All_Total – All_MA	Derived
All_FF_XMA	All Fully Funded Members Excluding Medicare Advantage Members	All Fully Funded Members, excluding MA members: ALL_FullyFunded – All_MA	Derived
All_PDP	All Prescription Drug Plan	Total number of members who have the Medicare Prescription Drug Plan coverage, where the health plan associated with the Medicare contract assumes financial risk	Membership
Mkt_Shr_Tot	Market Share Total	All members for a health plan as a percentage of total members for all health plans in the states where the health plan operates: All_Total / Mkt_Total	Derived
Mkt_Shr_FF	Market Share Fully Funded	All fully funded members for a health plan as a percentage of fully funded members for all health plans in the states where the health plan operates: All_FullyFunded / Mkt_FF	Derived
Mkt_Shr_MA	Market Share Medicare Advantage	All Medicare Advantage members for a health plan as a percentage of total Medicare Advantage members for all health plans in the states where the health plan operates: All_Total / Mkt_Total	Derived
Mkt_Shr_Tot_XMA	Market Share Total Excluding Medicare Advantage	All members excluding MA members for a health plan as a percentage of all members excluding MA members for all health plans in the states where the health plan operates: All_Total_XMA / Mkt_Tot_XMA	Derived
Mkt_Shr_FF_XMA	Market Share Fully Funded Excluding Medicare Advantage	All fully funded members excluding MA members for a health plan as a percentage of all fully funded members excluding MA members for all health plans in the states where the health plan operates: All_FF_XMA / Mkt_FF_XMA	Derived
Mkt_Shr_MA_Bin	Market Share Medicare Advantage, Binned	Market share of MA product (Mkt_Shr_MA), binned:	Derived

		• "<1%" – under 1 percent • "1 to 5%" – from 1 to 5 percent • "5 to 20%" – from 5 to 20 percent • ">20%" – greater than 20 percent	
Mkt_Shr_Tot_Bin	Market Share Total, Binned	Market share of all products (Mkt_Shr_Tot), binned: • "<1%" – under 1 percent • "1 to 5%" – from 1 to 5 percent • "5 to 10%" – from 5 to 10 percent • "10 to 20%" – from 10 to 20 percent • ">20%" – greater than 20 percent	Derived
ST_Total	State Total Membership	Total membership for a health plan in only those states where the plan provides the specific Medicare Advantage contract	Membership
ST_FF	State Fully Funded Membership	Membership in all fully funded products in only those states where the plan provides the specific Medicare Advantage contract	Membership
Corporate Strategy: Business Focus			
All_Perc_MA	Percentage of Members in Medicare Advantage	Total Medicare Advantage members of a health plan as a percentage of its total members; reflects its level of focus on Medicare Advantage products versus other products: All_MA / All_Total	Derived
All_Perc_FF	Percentage of Members in Fully Funded products	Total fully funded members of a health plan as a percentage of its total members; reflects its level of focus on managed care products versus other products: All_FullyFunded / All_Total	Derived
Corporate Strategy: Build Core Capabilities			
Stars_Calc_2019	Calculated CMS Stars Score for 2019	Weighted average of individual quality metric scores for 2019. Individual metrics and their weights are provided by CMS.	Calculated based on CMS data
Stars_Calc_2018	Calculated CMS Stars Score for 2018	Weighted average of individual quality metric scores for 2018. Individual metrics and their weights are provided by CMS.	Calculated based on CMS data
Stars_Calc_2017	Calculated CMS Stars Score for 2017	Weighted average of individual quality metric scores for 2017. Individual metrics and their weights are provided by CMS.	Calculated based on CMS data

Since many of the variables are derived from other variables (e.g., market share),

collinearity is a risk during modeling. The final model does not have any collinear variables. As

is done with the dependent variable, symbox and ggdensity functions were run on the

membership variables to evaluate the distribution of the data as well as the data symmetry

using different power transformations. For example, this is done on the variable Total_MA which

is defined as the total number of Medicare Advantage members in the health plan. The output is

shown on the figures below.

Figure 6: Symbox Transformation of Total_MA (Total Number of Medicare Advantage members in the health plan)
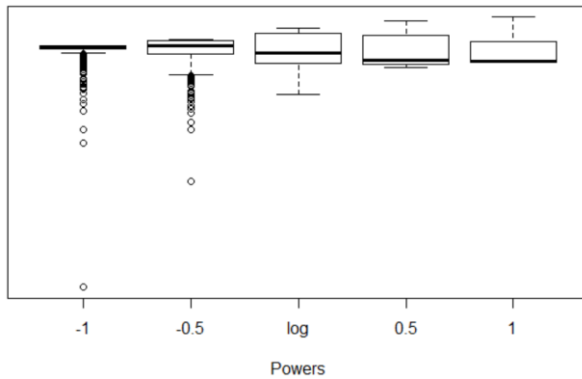
Figure 7: GG Density Plot for Total_MA (Total Number of Medicare Advantage members in the health plan)
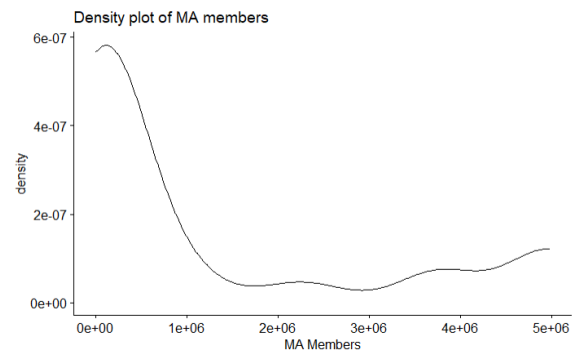
Far from being normally shaped, the distribution shows that a significant larger proportion of plans with fewer than 1 million members, smallest proportion of plans in the middle, and a larger proportion of plans with members at the right side of the distribution. Based on the symbox transformation output, the log transformation is used to create Total_MA1. The symbox output and GG density plot is displayed in Figures 8 and 9.

Figure 8: Symbox Transformation of Total_MA1

Figure 9: GG Density Plot for Total_MA1

The other measures of membership appear to have this feature, a non-normal distribution of membership with higher proportions of plans at the high (rightmost) and low (leftmost) ends. As

a result, these other measures go through a similar process of evaluation to determine the best transformations for use by the model.

An analysis is performed to determine correlation between the various independent variables against the response to determine which variables are most correlated with the response variabl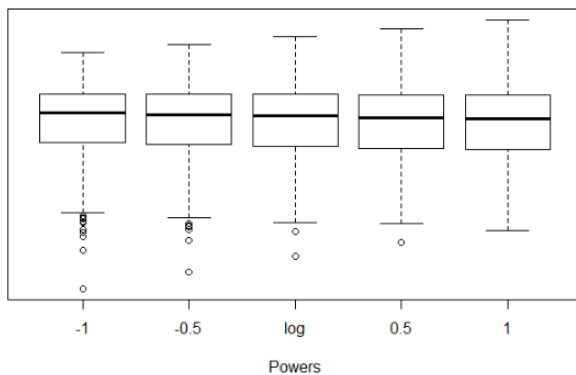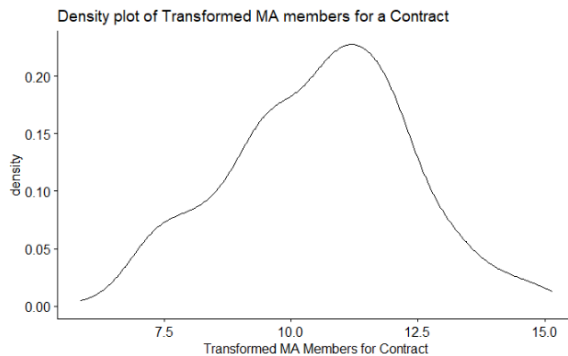e. Given the non-normal nature of the distribution, the Kendall method is selected over the Pearson method. The results of the top 20 variables, sorted in descending order based on the absolute value of tau is shown below in Table 4.

Table 4: Top 20 Variables with Highest Coefficient of Correlation with CMS Scores (Dependent Variable)

#	Variable	Tau	Absolute Value of Tau
1	Stars_Calc_2019	0.5937	0.5937
2	Stars_Calc_2018	0.5456	0.5456
3	Stars_Calc_2017	0.5390	0.5390
4	Contract_MA1	0.2105	0.2105
5	Mkt_Shr_MA2	0.1730	0.1730
6	All_Perc_FF1	(0.1072)	0.1072
7	All_Perc_MA1	0.0971	0.0971
8	DisasterPerc2017	0.0822	0.0822
9	Mkt_Shr_Tot2	0.0594	0.0594
10	Mkt_Tot_XMA2	(0.0560)	0.0560
11	Mkt_FF_XMA2	(0.0550)	0.0550
12	Mkt_Tot2	(0.0497)	0.0497
13	Mkt_MA1	(0.0492)	0.0492
14	Mkt_FF1	(0.0481)	0.0481
15	FF_Perc_Mkt2	(0.0465)	0.0465
16	All_MA1	0.0462	0.0462
17	DisasterPerc2018	0.0423	0.0423
18	All_FF_XMA1	(0.0409)	0.0409
19	Mkt_Shr_Tot_XMA	0.0406	0.0406
20	Mkt_Shr_FF2	0.0405	0.0405

CHAPTER 3
Methodology

The overall approach uses linear regression beginning with stepwise regression (stepAIC in R) to develop a candidate model as a starting point. The primary goal of this study is to identify and evaluate potential relationships of key drivers on the CMS star scores. So, from the models created by the stepwise regression, further analysis and modeling is performed iteratively using both BIC (Bayesian Information Criterion) and AIC (Akaike Information Criterion) to evaluate the model, VIF (Variable Inflation Factor) to evaluate the presence of multicollinearity, and Cook's Distance to identify and select out large outliers which skew the model results.

3.1 Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC)

The Akaike Information Criterion (AIC) is one of the estimators for model selection. AIC is an estimator of prediction error and its value can be compared to the value of other models to compare the quality of the model against others. AIC [4] is defined as:

$$AIC = 2k - 2ln(L)$$

In the formula above, k is the number of estimated parameters in the model, and L is the maximum value of the likelihood function for the model. So, AIC rewards goodness of fit while assessing a penalty as the number of estimated parameters increases. Since the user is selecting models with the lower AIC value, the penalty component of the formula (2k) discourages overfitting by adding to the score as the number of parameters is increased in the model.

Similarly, the Bayesian Information Criterion (BIC) is used to evaluate models during model selection. It is similar to the AIC but assigns a greater penalty for models with a large number of parameters than AIC. BIC [5] is defined as:

$$BIC = kln(n) - 2ln(L)$$

In the formula above, k represents the number of parameters in the model, n is the number of data points, and L is the maximized value of the likelihood function of the model. As a result, BIC results in a significantly more parsimonious model which is discussed on the results section of this thesis.

3.2 Variance Inflation Factor

To identify multicollinearity, where two or more predictor variables are closely related to each other, the thesis relies on the Variance Inflation Factor (VIF). The minimum value of VIF is 1 if there is no collinearity among the predictors. The VIF [6] is defined as:

$$VIF(\hat{\beta}_j) = \frac{1}{1 - R^2_{X_j | X_{-j}}}$$

In the formula, $R^2_{X_j | X_{-j}}$ is the R^2 value from a regression of X_j against all other predictor variables. If the VIF is 10, multicollinearity is high. For this study, if VIF values are above 5, this is considered the cutoff for elimination.

3.3 Cook's Distance
To identify data points with large outliers or high leverage that may disproportionately distort and affect the accuracy of the regression model, we use Cook's distance. Cook's distance [7] is defined as:

$$D_i = [\frac{1}{k}]t_i^2 \frac{h_{ii}}{1 - h_{ii}}$$

14

Where k is the number of covariates or predictors, t_i is the internally studentized residual, and h_{ii} is leverage for each observation. As the studentized residual becomes larger and as the leverage becomes larger and closer to 1, Cook's distance value increases.

CHAPTER 4
Results

4.1 Stepwise Regression

The method for the thesis uses stepwise regression function in R using both BIC and AIC as determinants for creating the starting model. Although AIC penalizes models that have greater number of parameters, since it rewards for predictive power, it has the risk of having too many terms in the model, as evidenced in this case. Using BIC creates a more parsimonious model. The figure below provides a summary of the initial models using stepwise regression and AIC vs BIC as the determinants for model selection.

Figure 10: Summary of Outputs for Initial Models from Stepwise Regression Using AIC and BIC

Parameter Used	Number of Predictors	AIC Value	BIC Value	Adjusted R-squared
AIC	14	-79.4311	-19.5944	0.7136
BIC	7	-80.4957	-46.8375	0.7083

Interestingly, although the AIC model has more predictors than the BIC model (14 vs 7), the model created by the BIC method has better AIC and BIC values than the model created by the AIC method. The model outputs are in Appendices B - E. Both initial models have at least one predictor that has a high degree of multicollinearity requiring additional model pruning. The adjusted R-squared for the models with 14 and 7 predictors are 0.7136 and 0.7083 respectively with a difference of 0.0053 for the adjusted R-squared measure. An ANOVA test comparing the two models results in a p-value of 0.08771, which is greater than the 0.05 value. As a result, we

conclude that the AIC model does not provide a statistically significant better fit to the data than does the BIC model. The results of the ANOVA test are in Appendix F.

4.2 Addressing Multicollinearity

The next step is to identify predictors that have high levels of multicollinearity with other predictors and rerun the models. The approach is to remove the predictor with the highest VIF value, then run the model to calculate and review the new VIF values. This iterative process stops when none of the VIF values are greater than five. At the end of this step, the new model created by the BIC approach has six predictor variables and the model created by the AIC approach has eight parameters. However, in the AIC model, three of the parameters have p-values that are not statistically significant and are removed. The table below shows the final predictors in BIC and AIC model.

Table 5: Summary of Predictors from Final BIC and AIC Models

Predictor	BIC Model	AIC Model
Stars_Calc_2019	X	X
Stars_Calc_2018	X	X
Stars_Calc_2017	X	X
Mkt_FF_XMA2	X	
All_Perc_MA1	X	X
All_Perc_FF1	X	X

For the final model, the BIC model was chosen over the AIC model because it had lower values for both the BIC and AIC measure. The BIC value for the BIC model was -44.7214 vs -44.1952 for the AIC model, and the AIC value for the BIC model was -74.6397 vs -70.3738. This final model is discussed in section 4.5 below.

4.3 Cook's Distance

Figure 11 shows the results of the plot function, option 4, for the final BIC model. This figure shows that none of the records have Cook's distance value of 0.05, which would be a consideration for removal, so there is no need to remove any records.

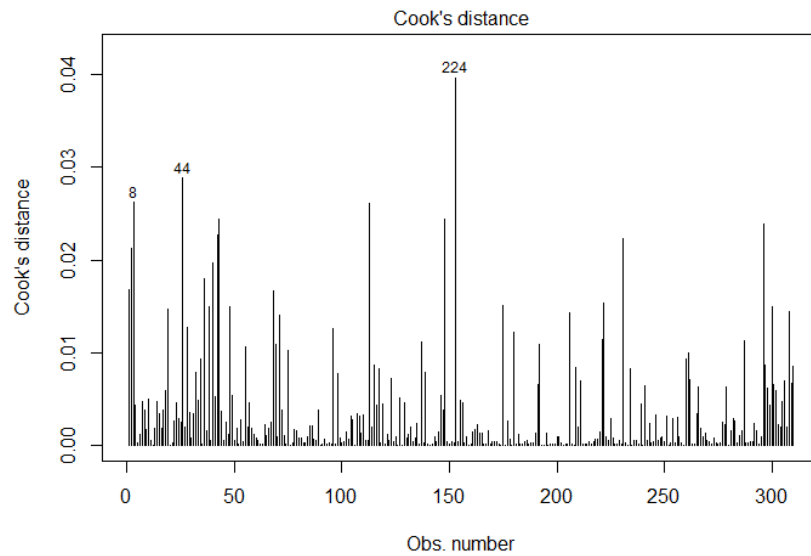**Figure 11: Cook's Distance Plot of the Final BIC Model**

Figure 12: Cook's Distance Plot of the Final AIC Model
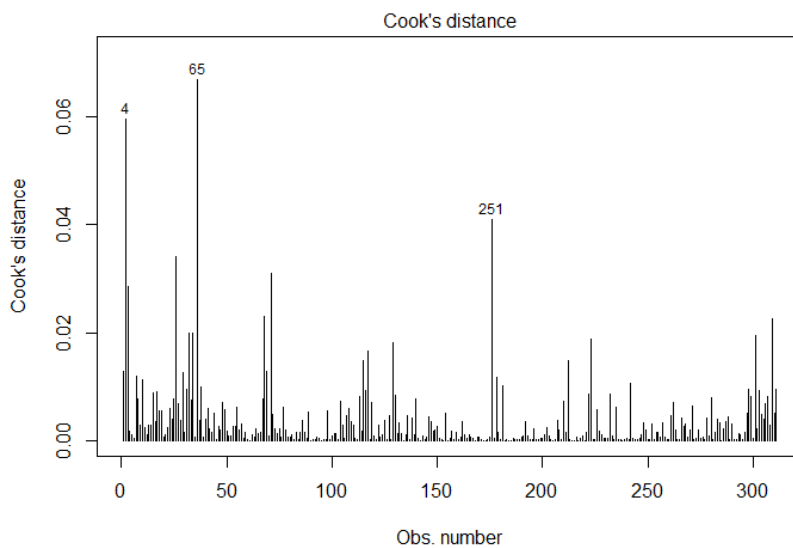
Figure 12 above shows the output with Cook's distance for the final model using AIC method. When records 4 and 65 are removed and rerun, this did not materially change the model. Another way to identify leverage points that may influence findings is to use the influencePlot function in R, which creates a bubble chart of the Studentized residuals against hat values. The circles represent the observations and the size of the circle represent the Cook's distance value. Figure 13 below shows the output of the influencePlot, identifying in addition to observations 4, 8, and 65, points 83 and 407. The Cook's distance value of 407 is roughly the same size as for observation 8 and the Cook's distance value of observation 83 is very small.

Figure 13: Output of R InfluencePlot
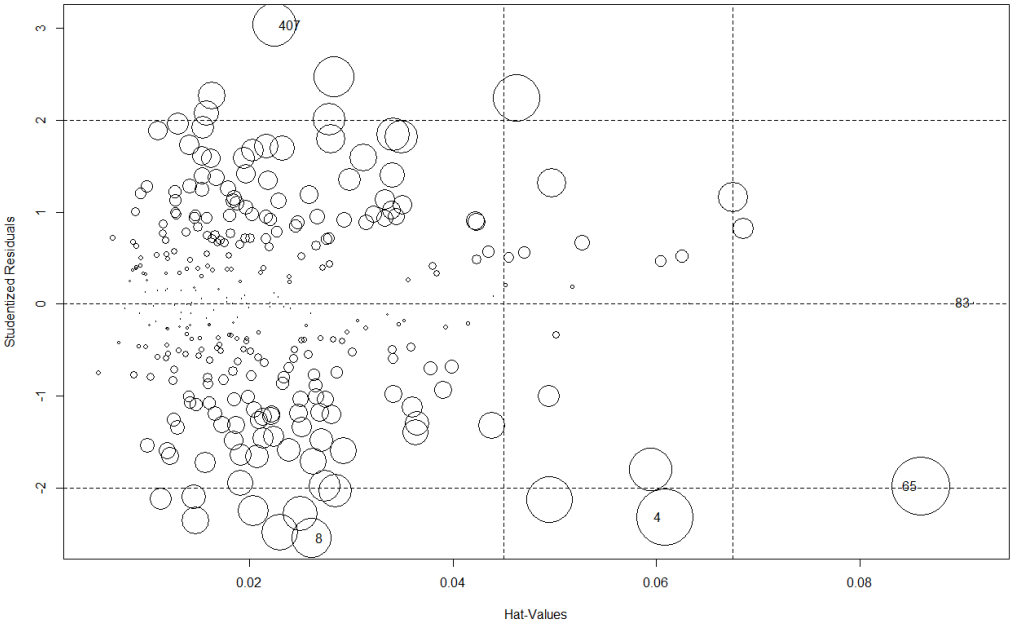
4.4 Final Model

The final model is based on the final model that comes from the stepwise regression method using BIC as the criterion. The output and VIF values of the final model are shown below. Its BIC value is -44.7214 and AIC value is -74.6397.

Figure 14: Final Model Outputs and AIC/BIC Values

Parameter	Description	Estimate	Std Error	t-value	Pr(>|t|)
Intercept	Intercept	1.317	0.1723	7.644	2.78e-13
Stars_Calc_2019	2019 CMS Star Score	0.3649	0.04240	8.606	4.12e-16
Stars_Calc_2018	2018 CMS Star Score	0.1861	0.04667	3.988	8.35e-6
Stars_Calc_2017	2017 CMS Star Score	0.2320	0.04301	5.393	1.40e-7
Mkt_FF_XMA2	Market Size, Fully Funded Products Excluding Medicare Advantage	-7.128e-6	2.866e-6	-2.487	0.01341
All_Perc_MA1	Health Plan's Medicare Advantage Members Aa % of Its Total Members	0.0312	0.01118	2.792	0.00557
All_Perc_FF1	Health Plan's Medicare Fully Funded Members Aa % of Its Total Members	-0.3828	0.08903	-4.299	2.31e-6

Notes:

Residual standard error: 0.2115 on 304 degrees of freedom (98 observations deleted due to missingness)

Multiple R-squared: 7076, Adjusted R-squared: 0.7018

F-statistic: 122.6 on 6 and 304 DF, p-value: < 2.2e-16

AIC: -74.6397

BIC: -44.7214

Figure 15: VIF Values for Final Model Parameters

Stars_Calc_2019	Stars_Calc_2018	Stars_Calc_2017	Mkt_FF_XMA2	All_Perc_MA1	All_Perc_FF1
2.371843	2.591184	2.532971	1.445209	1.253834	1.654487

When comparing the final model against the list of variables in Table 4 with coefficients of correlation, it is noteworthy that five out of six variables in the final model are in the top ten variables with the highest absolute values in Table 4. The only variable, Mkt_FF_XMA2, outside of this list is ranked eleventh place. From Table 4, Contract_MA1, Mkt_Shr_MA2, DisasterPerc2017 and Mkt_Shr_Tot2 are variables in the top 10, but these are not in the final model. Among these four variables, none of these variables are in the models produced stepwise regression using either AIC or BIC criteria.

Figures 16 to 18 provide additional information on the residuals created by the final model. Figures 16 and 18 show balanced residuals.

Figure 16: Residual vs Fitted Plot of Final Model

Residuals vs Fitted
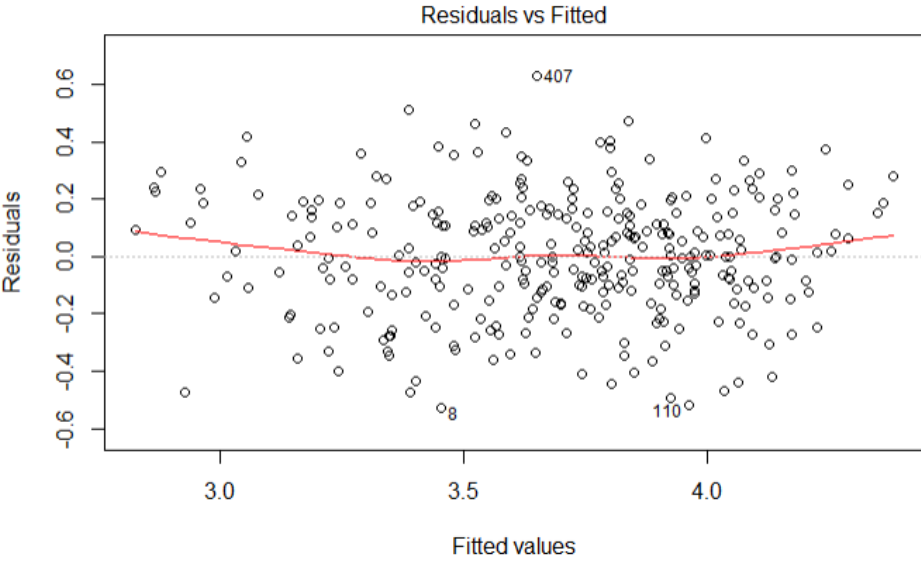
Figure 17: Normal Q-Q Plot of Final Model
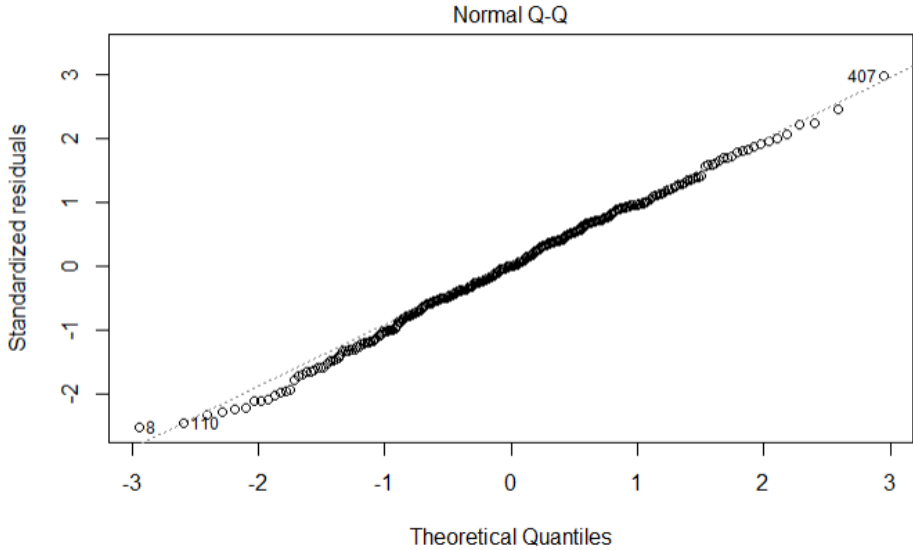
Normal Q-Q

Figure 17 shows that the residual follows a roughly normal distribution, which indicates that we can assume that the parameter estimates, significant tests, and confidence intervals are valid.

Figure 18: Scale-Location Plot of Final Model
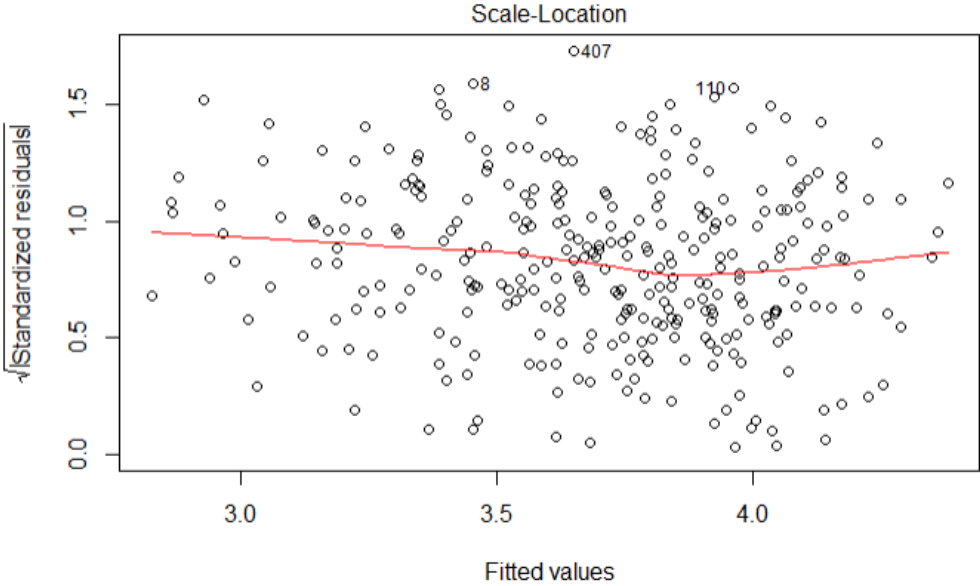
Figure 18 shows that the average magnitude of the standardized residuals does not change greatly as a function of the fitted values (red line is roughly horizontal) and the overall variability (spread) around the red line does not change much as a function of the fitted line.

CHAPTER 5
Discussion

5.1 Discussion of Model

Overall, the final model has 6 parameters not including the intercept and an Adjusted R-squared value of 0.7018. The BIC value is -44.7214 and AIC value is -74.6397. The final model presents interesting findings for what is excluded as well as included. It is noteworthy that factors associated with the health plans' organizational types are not selected in the final model. For example, the type of administrative organization, its profit versus nonprofit status, whether it is a private or publicly traded company does not factor in the final model. In addition, Special Needs Plans (SNP), health plans which have additional expertise providing enhanced services for frail Medicare members, do not appear to be any better than other plans in achieving high CMS scores.

Not surprisingly, the 2020 CMS Star Score for a plan is strongly associated with the past performance in 2019, 2018, and 2017 and the parameter estimates. This suggests that the higher a plan's CMS score is in past years, the higher these scores are in 2020. Put another way, health plans can build capabilities to improve its CMS score, and this capability carries forward into future years. As expected, the parameter estimate for the CMS score in 2019 is higher than it is for 2018 and 2017. However, interestingly, the data shows that the parameter estimate for 2017 is higher than for 2018, indicating that its 2020 score has more in common to its 2017 performance than its 2018 performance.

Industry structure plays a role in the CMS score, according to the model. Mkt_FF_XMA2 is the total number of fully funded members excluding Medicare Advantage (MA) for all health

plans in all of the states where the plan does business, raised to the 0.5 power. It is a measure of how large the industry is as it pertains to fully funded members excluding MA. As a reminder, fully funded relates to cases where the health insurance plans and its physician and hospital networks share financial risk for the costs associated with patient care. Mkt_FF_XMA2 measures scale of a plan's managed care business excluding MA.

The model also has two parameters which indicate that CMS scores may be a result of its business focus, All_Perc_MA1 and All_Perc_FF1. Respectively, these are the log transformations of the percentage of a health plan's members that are in Medicare Advantage and fully funded plans. The parameter estimate for All_Perc_MA1 is 0.03122, indicating that as the percentage of a health plan's membership, and hence its business focus in Medicare Advantage, increases so does its CMS score. Meanwhile, the value of All_Perc_FF1 is negative (-0.3828), indicating that health plans that focus more on its overall managed care business generally are associated with poorer performance on its CMS score.

The parameter estimate of Mkt_FF_XMA2 is -0.000007, indicating that as the total market size of fully funded membership (excluding MA) grows, this has a negative effect on the CMS score. This is counterintuitive at first. Medicare Advantage is a subset of fully funded insurance schemes. One could argue, shouldn't large markets for fully funded insurance provide insurance companies greater scale and the ability to provide better managed services that improve CMS scores? One possible explanation is that there are other unmeasured factors about these larger managed care markets that are why these markets are not performing as well. While this factor is counterintuitive, Mkt_FF_XMA2 does not appear to have a very strong influence on the model. Rerunning the model after removing this parameter yields a model with an adjusted R-squared value of 0.6968, which is very close to the final model's adjusted R-squared value of 0.7018.

Another possible explanation is that many of these companies with high percentage of fully funded members are those which do not specialize in the Medicare or Medicare Advantage space. As mentioned above, Medicare Advantage, which is a fully funded Medicare plan, represents only 34% of Medicare, so companies that focus on the Medicare Advantage and Medicare space likely have lower percentage of fully funded members, whereas there are many large national insurance companies that focus more on other markets which have significantly higher percentage of members in fully funded plans.

CHAPTER 6
Conclusion

6.1 Conclusions

From the modeling, there are some initial potential findings for health plans. For one, organizational attributes (e.g., profit vs non-profit and public vs. private) do not appear to have a statistically significant association with the CMS score. What appears to positively affect CMS scores are having higher CMS scores in prior years, indicating past competencies or investments in capabilities improve CMS scores. Another factor that raises CMS scores is a business focus on its Medicare Advantage (MA) plans, in other words having a greater share of its members be on Medicare Advantage plans. Meanwhile, plans that have a higher percentage of its members in its fully funded managed care plans tend to be associated with lower scores.

Industry structure seems to have an effect in the model. Mkt_FF_XMA2 reflects the total size of the managed care market excluding MA. This suggests that in larger managed care markets, plans have lower CMS scores. One possible explanation could be that in these larger managed care markets, health plans are less focused on the MA product. Another possibility is that in these larger markets, there is less of a close relationship between plans and their providers, which is required to better manage members and obtain high CMS scores.

For health plans interested in improving the CMS scores, the potential strategic implications from the model are that they should invest to improve their capabilities to improve their CMS scores, which carry forward and have significant multiyear effects. In fact, their financial investment models measuring return on investment (ROI) should include estimates with returns in subsequent years. The ROI inputs for these subsequent years can be produced

from a statistical model. Another implication is that in terms of M&A activity, achieving higher market share and achieving scale is not associated with obtaining higher CMS scores. This is noteworthy since many health plans including those which specialize in Medicare Advantage have merged precisely to increase scale and achieve higher market share.

Finally, the study which uses statistics is a test case for determining the feasibility of using advanced analytics such as statistics and machine learning to provide insights and better inform health plans so they can make well-informed strategic and operation decisions based on data and using quantitative methods and frameworks. The Medicare Advantage program is very large, accounting for a third of a trillion dollars a year on spending. Even minor improvements in the program would have significant benefits for health plans, providers, consumers, and society at large. Expanding this out further, healthcare expenses represent 17.7% of US GDP [8] and there are huge inefficiencies in the US healthcare system. Health plans can significantly overcome these inefficiencies and reduce overall costs by making more optimal spending and resource allocation decisions by leveraging statistics and machine learning.

6.2 Limitations

One of the key limitations of this study is that it is a one-year snapshot and is limited by the fact that there is only one year of membership data. As a result, it is not known how membership changes over time, particularly in its Medicare Advantage membership, may affect its CMS score. Despite its limitations, the model does provide some valuable possible relationships that could be tested in future studies as well as providing health plans with an industrywide perspective on what appears to affect CMS Star score performance.

6.3 Future Analysis

Future analysis should include multiple years of membership data, which would allow the model to evaluate the association between membership change for health plans and its CMS

star scores. Given the importance of CMS scores in past years on current year CMS scores, future studies can further examine and refine the relationship between past CMS score performance and subsequent year performance. To do this analysis, the study can use other methodologies such as match analysis, using past year CMS score achieved (for example, achieving a score of 4 or more stars) as a "treatment" variable and determining this effect in future years. The results of this analysis could provide greater support for establishing causality (becoming a high achiever in CMS star scores will help plans become high achievers in subsequent years) rather than establishing mere association. Match analysis can be used with other variables in the final model as well by assigning these other variables as the "treatment effect".

More time series data would be able to examine answer other questions as well such as whether there are secular trends that influence CMS star scores? For example, do some plans (e.g., large plans) generally do better over time? Based on this study, the expectation would be that they do not since there do not appear to be any variables capturing scale in the model. However, the trend could be different. Beyond scale advantages, more time series data could be used to study additional temporal factors such as socioeconomic data and the influence of major events such as COVID-19 on CMS scores.

Finally, future studies could also include data on local provider conditions on CMS scores. One possible hypothesis could be that areas where there are fewer or a single major provider such as a local hospital could result in lower CMS scores for plans in these areas. In these areas, providers have significant power over plans in terms of reimbursement rates and may be less incentivized by the CMS bonus sharing offered by health plans. On the other area of the spectrum, areas where there are significantly fragmented provider markets, providers may be disincentivized to participate fully in CMS plans because they lack the scale, and the bonus incentives are not large enough to encourage their full attention.

Appendixes

Appendix A: Individual CMS Measures and Weights, Year 2020 (1 of 3)

Measure ID	Measure Name	Description	Weighting Category	Score Weight
C01	Breast Cancer Screening	Percent of female plan members aged 52-74 who had a mammogram during the past 2 years.	Process Measure	1
C02	Colorectal Cancer Screening	Percent of plan members aged 50-75 who had appropriate screening for colon cancer.	Process Measure	1
C03	Annual Flu Vaccine	The percentage of sampled Medicare enrollees (denominator) who received an influenza vaccination (numerator).	Process Measure	1
C04	Improving or Maintaining Physical Health	Percent of plan members whose physical health was the same or better than expected after two years.	Outcome Measure	3
C05	Improving or Maintaining Mental Health	Percent of plan members whose mental health was the same or better than expected after two years.	Outcome Measure	3
C06	Monitoring Physical Activity	Percent of senior plan members who discussed exercise with their doctor and were advised to start, increase, or maintain their physical activity during the year.	Process Measure	1
C07	Adult BMI Assessment	Percent of plan members with an outpatient visit who had their Body Mass Index (BMI) calculated from their height and weight and recorded in their medical record.	Process Measure	1
C08	Special Needs Plan (SNP) Care Management	Percent of members whose plan did an assessment of their health needs and risks in the past year. The results of this review are used to help the member get the care they need.	Process Measure	1
C09	Care for Older Adults – Medication Review	Percent of plan members whose doctor or clinical pharmacist reviewed a list of everything they take (prescription and non-prescription drugs, vitamins, herbal remedies, other supplements) at least once a year.	Process Measure	1
C10	Care for Older Adults – Functional Status Assessment	Percent of plan members whose doctor has done a functional status assessment to see how well they are able to do Activities of Daily Living such as dressing, eating, and bathing.	Process Measure	1
C11	Care for Older Adults – Pain Assessment	Percent of plan members who had a pain screening at least once during the year.	Process Measure	1
C12	Osteoporosis Management in Women who had a Fracture	Percent of female plan members who broke a bone and got screening or treatment for osteoporosis within 6 months.	Process Measure	1
C13	Diabetes Care – Eye Exam	Percent of plan members with diabetes who had an eye exam to check for damage from diabetes during the year.	Process Measure	1
C14	Diabetes Care – Kidney Disease Monitoring	Percent of plan members with diabetes who had a kidney function test during the year.	Process Measure	1
C15	Diabetes Care – Blood Sugar Controlled	Percent of plan members with diabetes who had an A1C lab test during the year that showed their average blood sugar is under control.	Intermediate Outcome Measure	3
C16	Rheumatoid Arthritis Management	Percent of plan members with rheumatoid arthritis who got one or more prescriptions for an anti-rheumatic drug.	Process Measure	1

Measure ID	Measure Name	Description	Weighting Category	Score Weight
C17	Reducing the Risk of Falling	Percent of plan members with a problem falling, walking, or balancing who discussed it with their doctor and received a recommendation for how to prevent falls during the year.	Process Measure	1
C18	Improving Bladder Control	Percent of plan members with a urine leakage problem in the past 6 months who discussed treatment options with a provider.	Process Measure	1
C19	Medication Reconciliation Post-Discharge	This shows the percent of plan members whose medication records were updated within 30 days after leaving the hospital. To update the record, a doctor or other health care professional looks at the new medications prescribed in the hospital and compares them with the other medications the patient takes. Updating medication records can help to prevent errors that can occur when medications are changed.	Process Measure	1
C20	Plan All-Cause Readmissions	Percent of senior plan members discharged from a hospital stay who were readmitted to a hospital within 30 days, either for the same condition as their recent hospital stay or for a different reason.	Outcome Measure	3
C21	Statin Therapy for Patients with Cardiovascular Disease	This rating is based on the percent of plan members with heart disease who get the right type of cholesterol-lowering drugs. Health plans can help make sure their members are prescribed medications that are more effective for them.	Process Measure	1
C22	Getting Needed Care	Percent of the best possible score the plan earned on how easy it is for members to get needed care, including care from specialists.	Patients' Experience and Complaints Measure	1.5
C23	Getting Appointments and Care Quickly	Percent of the best possible score the plan earned on how quickly members get appointments and care.	Patients' Experience and Complaints Measure	1.5
C24	Customer Service	Percent of the best possible score the plan earned on how easy it is for members to get information and help from the plan when needed.	Patients' Experience and Complaints Measure	1.5
C25	Rating of Health Care Quality	Percent of the best possible score the plan earned from members who rated the quality of the health care they received.	Patients' Experience and Complaints Measure	1.5
C26	Rating of Health Plan	Percent of the best possible score the plan earned from members who rated the health plan.	Patients' Experience and Complaints Measure	1.5
C27	Care Coordination	Percent of the best possible score the plan earned on how well the plan coordinates members' care. (This includes whether doctors had the records and information they needed about members' care and how quickly members got their test results.)	Patients' Experience and Complaints Measure	1.5
C28	Complaints about the Health Plan	Percent of members filing complaints with Medicare about the health plan.	Patients' Experience and Complaints Measure	1.5
C29	Members Choosing to Leave the Plan	Percent of plan members who chose to leave the plan.	Patients' Experience and Complaints Measure	1.5

Measure ID	Measure Name	Description	Weighting Category	Score Weight
C30	Health Plan Quality Improvement	This shows how much the health plan's performance improved or declined from one year to the next. If a plan receives 1 or 2 stars, it means, on average, the plan's scores declined (got worse). If a plan receives 3 stars, it means, on average, the plan's scores stayed about the same. If a plan receives 4 or 5 stars, it means, on average, the plan's scores improved.	Improvement Measure	5
C31	Plan Makes Timely Decisions about Appeals	Percent of plan members who got a timely response when they made an appeal request to the health plan about a decision to refuse payment or coverage.	Measures Capturing Access	1.5
C32	Reviewing Appeals Decisions	This rating shows how often an independent reviewer thought the health plan's decision to deny an appeal was fair. This includes appeals made by plan members and out-of-network providers. (This rating is not based on how often the plan denies appeals, but rather how fair the plan is when they deny an appeal.)	Measures Capturing Access	1.5
C33	Call Center – Foreign Language Interpreter and TTY Availability	Percent of time that TTY services and foreign language interpretation were available when needed by people who called the health plan's prospective enrollee customer service phone line.	Measures Capturing Access	1.5

Appendix B: Stepwise Regression Using BIC - Model Summary and AIC/BIC Values

```
Residuals:
     Min       1Q    Median        3Q       Max
-0.50930  -0.13049   0.00037   0.14388   0.64606

Coefficients:
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)      1.241e+00  1.726e-01   7.188 5.14e-12 ***
Stars_Calc_2019  3.912e-01  4.299e-02   9.100  < 2e-16 ***
Stars_Calc_2018  1.667e-01  4.669e-02   3.572 0.000413 ***
Stars_Calc_2017  2.217e-01  4.270e-02   5.192 3.82e-07 ***
All_Total2       3.244e-05  1.165e-05   2.784 0.005705 **
Mkt_FF_XMA2     -1.769e-05  4.735e-06  -3.736 0.000224 ***
All_Perc_MA1     3.586e-02  1.119e-02   3.206 0.001488 **
All_Perc_FF1    -2.622e-01  9.814e-02  -2.671 0.007964 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2092 on 303 degrees of freedom
  (98 observations deleted due to missingness)
Multiple R-squared:  0.7149,   Adjusted R-squared:  0.7083
F-statistic: 108.5 on 7 and 303 DF,  p-value: < 2.2e-16

AIC: -80.49567
BIC: -46.83753
```

Appendix C: Stepwise Regression Using BIC - VIF Summary

```
Stars_Calc_2019 Stars_Calc_2018 Stars_Calc_2017        All_Total2
       2.492174        2.650234        2.551936          5.186084

Mkt_FF_XMA2     All_Perc_MA1     All_Perc_FF1
   4.032729        1.282379         2.054975
```

Appendix D: Stepwise Regression Using AIC - Model Summary and AIC/BIC Values

```
Residuals:
     Min      1Q   Median      3Q      Max
-0.50557 -0.13535  0.00183  0.13065  0.62943

Coefficients:
                   Estimate Std. Error t value Pr(>|t|)
(Intercept)       1.639e+00  2.569e-01   6.379 6.84e-10 ***
DisasterPerc2018  1.586e-03  1.007e-03   1.576 0.116095
Mkt_Shr_Tot_XMA  -1.536e+00  6.239e-01  -2.462 0.014399 *
Mkt_Shr_FF_XMA    1.359e+00  6.553e-01   2.074 0.038989 *
Stars_Calc_2019   3.966e-01  4.367e-02   9.082  < 2e-16 ***
Stars_Calc_2018   1.657e-01  4.667e-02   3.550 0.000447 ***
Stars_Calc_2017   1.833e-01  4.430e-02   4.138 4.58e-05 ***
All_Total2        1.456e-03  7.911e-04   1.841 0.066631 .
All_FullyFunded2 -8.172e-04  4.528e-04  -1.805 0.072140 .
All_Total_XMA1   -1.432e-03  7.877e-04  -1.819 0.069980 .
All_FF_XMA1       8.439e-04  4.399e-04   1.918 0.056021 .
All_Perc_MA1      2.746e-02  1.935e-02   1.419 0.157027
All_Perc_FF1     -6.059e-01  1.935e-01  -3.132 0.001912 **
Mkt_Tot_XMA2      1.178e-04  5.854e-05   2.013 0.045070 *
Mkt_FF_XMA2      -1.738e-04  7.498e-05  -2.317 0.021169 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2073 on 296 degrees of freedom
  (98 observations deleted due to missingness)
Multiple R-squared:  0.7265,   Adjusted R-squared:  0.7136
F-statistic: 56.17 on 14 and 296 DF,  p-value: < 2.2e-16


AIC:  -79.4311
BIC:  -19.59442
```

Appendix E: Stepwise Regression Using AIC – VIF Summary

```
DisasterPerc2018  Mkt_Shr_Tot_XMA   Mkt_Shr_FF_XMA  Stars_Calc_2019
       1.147702        24.314976        18.720267         2.618506

Stars_Calc_2018  Stars_Calc_2017        All_Total2     All_FullyFunded2
2.697774         2.796632          24349.977421     3187.108113

All_Total_XMA1      All_FF_XMA1      All_Perc_MA1        All_Perc_FF1
22220.587667     2422.313764         3.908877            8.133840

Mkt_Tot_XMA2       Mkt_FF_XMA2
1034.392172      1030.050314
```

Appendix F: Results of ANOVA Test Comparing the AIC and BIC Models

```
Analysis of Variance Table

  Res.Df     RSS Df Sum of Sq       F  Pr(>F)
1    296 12.726
2    303 13.266 -7  -0.54046 1.7959 0.08771 .
```

References/Bibliography

[1] Kaiser Family Foundation. "An Overview of Medicare". https://www.kff.org/medicare/issue-brief/an-overview-of-medicare/

[2] MedPAC. "Health Care Spending and the Medicare Program." http://medpac.gov/docs/default-source/data-book/july2020_databook_entirereport_sec.pdf

[3] CMS website. https://www.cms.gov/Medicare/Prescription-Drug-Coverage/PrescriptionDrugCovGenIn/Downloads/2020-Star-Ratings-Data-Table-Oct-21-2019.zip

[4] Faraway, J.J. "Linear Models with R," Chapman & Hall, pp 90.

[5] Sheather, S. "A Modern Approach to Regression with R," Springer, pp 232-233.

[6] Ibid, pp 203.

[7] Faraway, J.J. "Linear Models with R," Chapman & Hall, pp 154.

[8] CMS website. https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/NationalHealthAccountsHistorical#:~:text=U.S.%20health%20care%20spending%20grew,spending%20accounted%20for%2017.7%20percent.