June 2020

Assessing Hospital Preparedness for COVID-19
Statistical Overview to Accompany Article

CONFIDENTIAL

JUNIPER ADVISORY

191 North Wacker Drive, Suite 900
Chicago, IL 60606
www.juniperadvisory.com

Jordan Shields | 312.506.3005 | jshields@juniperadvisory.com
Alex Normington | 312.506.3003 | anormington@juniperadvisory.com
Alex Voss | 312.506.3002 | avoss@juniperadvisory.com
Andrew Blank | 312.506.3004 | ablank@juniperadvisory.com

Page

1. Abstract

IMPACT OF SYSTEM STATUS ON SELECT HOSPITAL RESOURCES *abstract*

- System hospitals tend to have more ICU beds, higher case mix indices and are more likely to be a part of a Clinically Integrated Network or Accountable Care Organization than comparable standalone hospitals

- These attributes are particularly important findings during the COVID-19 pandemic, where ICU beds are at a premium, experience managing complex cases is essential and the ability to care for patients in appropriate settings is of utmost importance

- Findings:
  - **Critical care infrastructure**: Holding case mix index, CMS Hospital Compare score, discharges and length of stay constant, we found that, for example, a 100-bed community hospital that is part of a system is likely to have 1.5 more ICU beds than a 100-bed standalone hospital

  - **Preparedness for novel, complex cases**: A hospital that is part of a system will, on average, have a case mix index that is 0.05 higher than a similarly situated standalone hospital. For the typical hospital in the set with a CMI of 1.62, being in a system predicts a 3% increase. This indicates that systems work to grow their community hospitals by adding services and keeping patients close to home

  - **Ability to care for patients in appropriate settings**: With the model controlling for operating margin, Medicare payor mix and CMS Hospital Compare score, we found that being in a system has a statistically significant (positive) relationship with the odds of being a member of an ACO or CIN

2. Introduction

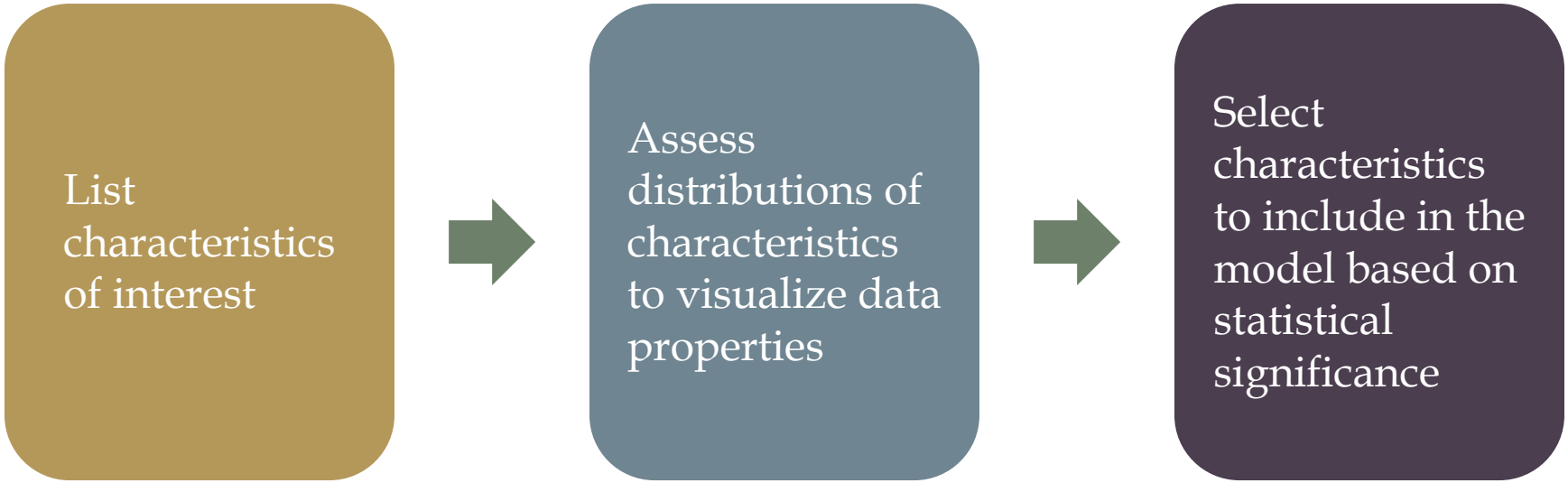
IMPACT OF SYSTEM STATUS ON SELECT HOSPITAL RESOURCES *research question*

- The COVID-19 pandemic highlighted certain resource gaps in U.S. healthcare
  - Critical care infrastructure
  - Preparedness for novel, complex cases
  - Ability to care for patients in appropriate settings

- We have used Medicare cost report data to better understand whether system hospitals are better positioned to fill these gaps than standalone facilities while correcting for confounding variables

- To test this, we used a dataset of characteristics of our defined population to create a model with coefficients that show the magnitude and direction of these relationships

- A simplified method for a linear model:



IMPACT OF SYSTEM STATUS ON SELECT HOSPITAL RESOURCES *research question*

- While we followed a similar approach for case mix index, we will describe our methodology for ICU beds here

- Is there a relationship between characteristics of hospitals in the US and the number of ICU beds they have?

- We examine several variables of interest:

  - **Hospital affiliation**
  - **Net patient revenue**
  - **Operating profit margin**
  - **EBITDA**
  - **Rural vs. non-rural**
  - **Case mix index**

  - **Number of discharges**
  - **Number of staffed beds**
  - **Adjusted patient days**
  - **Average length of stay**
  - **Payor mix**
  - **Hospital Compare score**

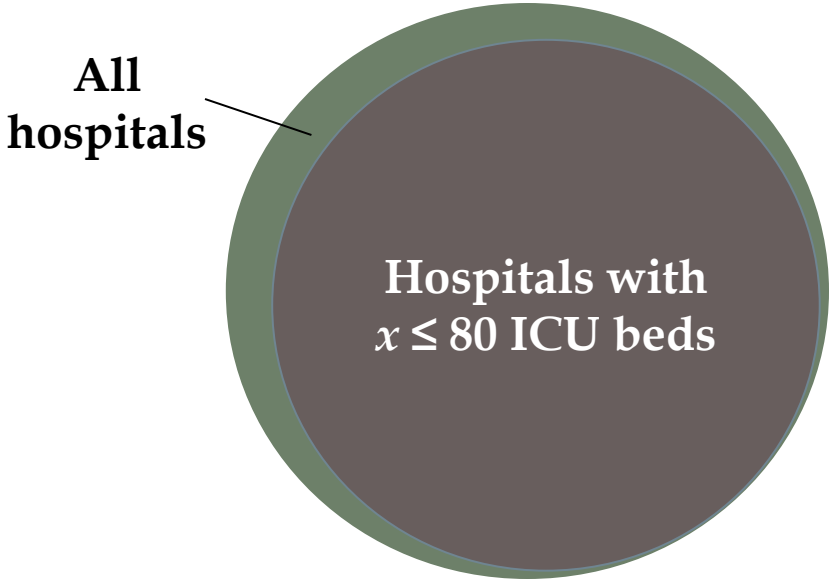
- The set of hospital data we examine is from CMS, collected in annual reports, made available by Definitive Healthcare in the spring of 2020
  - We can thereby take a view of the population rather than rely on a sample from the population

- Why develop a model in the first place?
  - A model can help reveal relationships between variables while controlling for other variables to give a 'truer' idea of an effect

IMPACT OF SYSTEM STATUS ON SELECT HOSPITAL RESOURCES *research question*

- First, we consider the set of all hospitals, examining:
  - Distributions of variables
  - One-to-one relationships
  - Linear model and coefficients
  - *See appendix for this model*

- Then, we focus on a subset of all hospitals defined as those having 80 or fewer ICU beds
- We made this distinction because:
  - There are very few standalone hospitals with >80 ICU beds
  - These very large hospitals are fundamentally different enterprises than community hospitals



3. ICU Beds

STATISTICAL DISTRIBUTION TYPES *data overview*

- In this analysis, the (continuous) data categories tend to be distributed in a few ways:

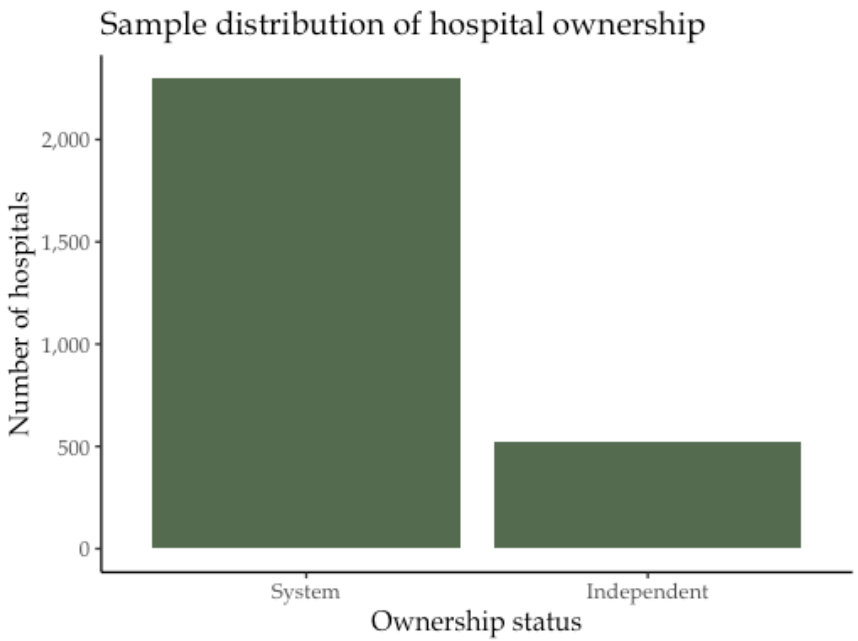
Normal

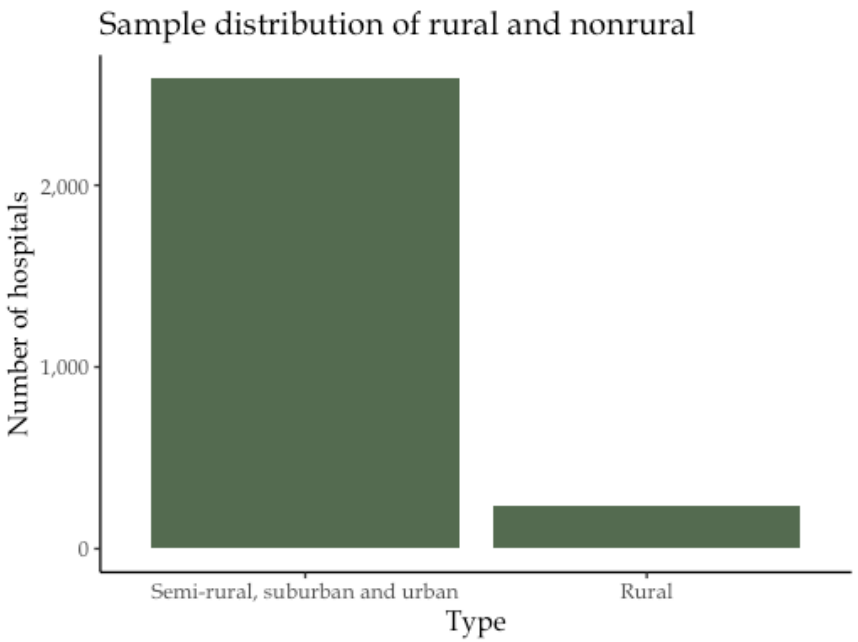
- Nice properties
- Symmetry
- Well-defined mean and variance

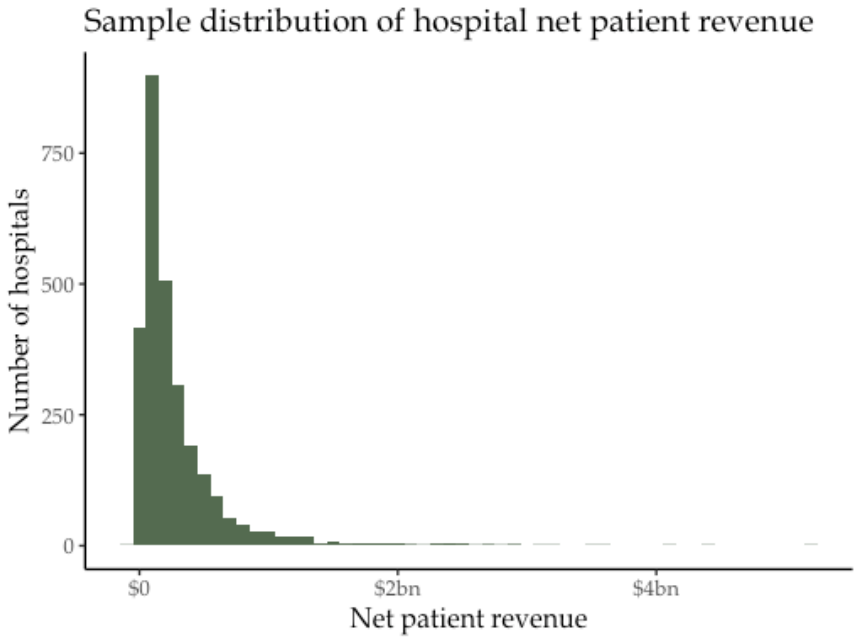
Power law

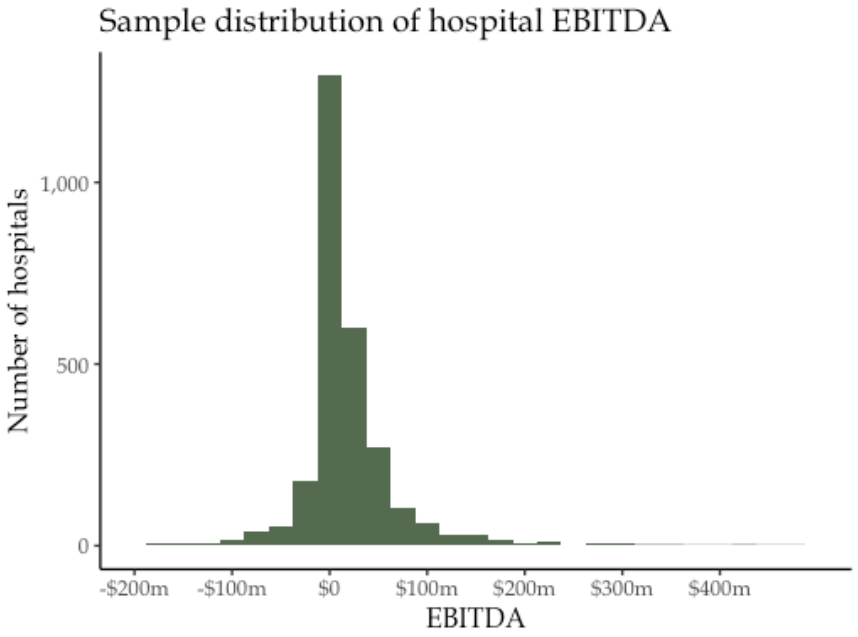
- Pareto principle, or the 80/20 rule
- Heavy skew, meaning some observations are far from the main cluster, or that the curve is asymmetric
- Possibly heavy or fat tail, meaning many observations are in the tail, especially compared to a normal distribution

VARIABLES OVERVIEW *distributions*

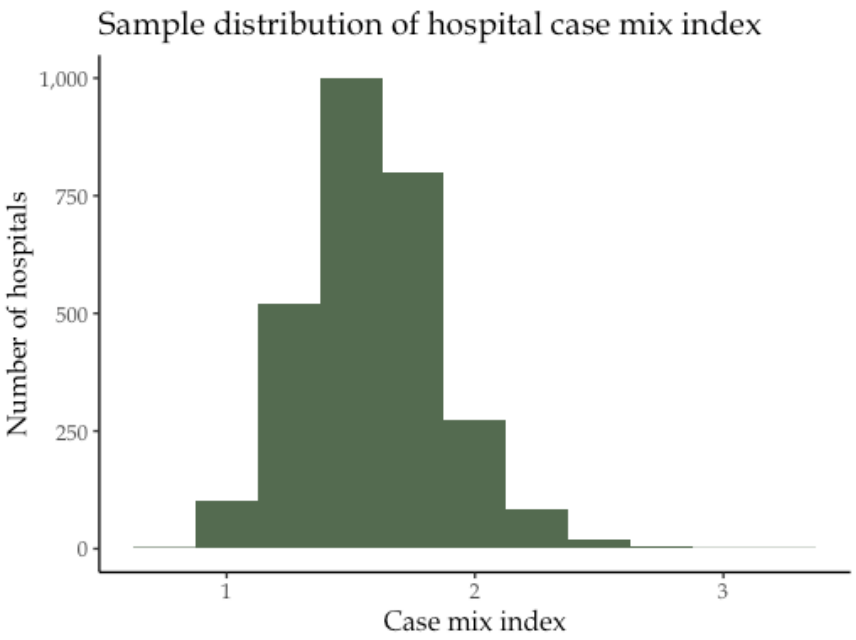


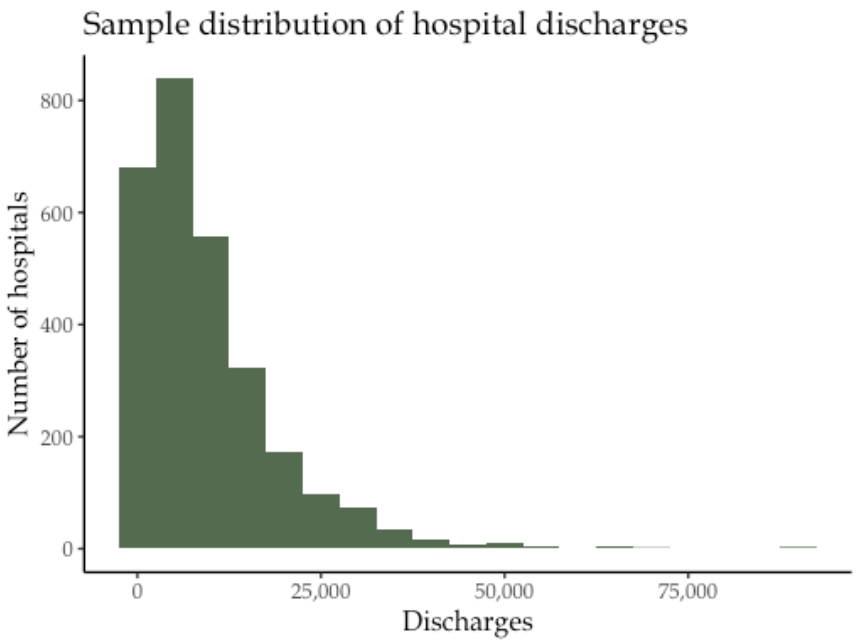


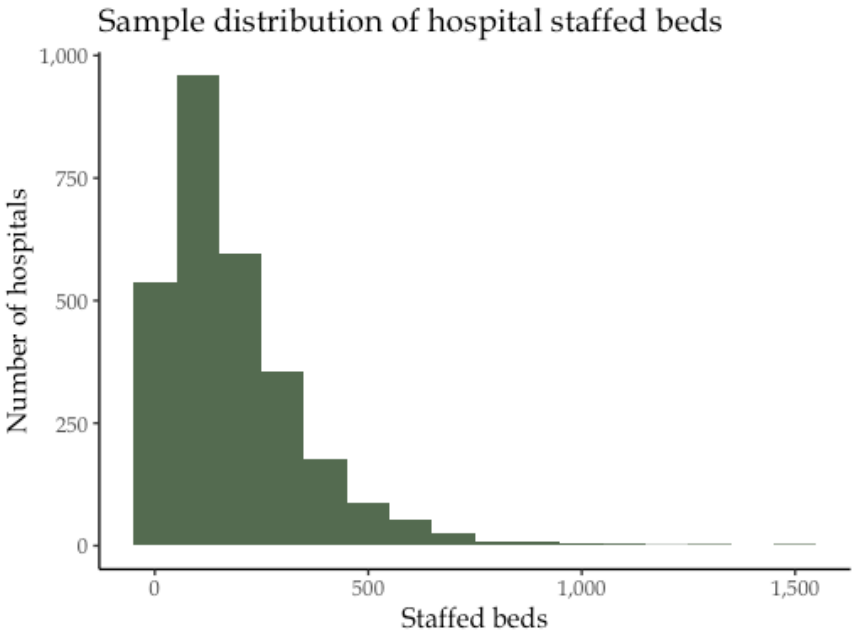


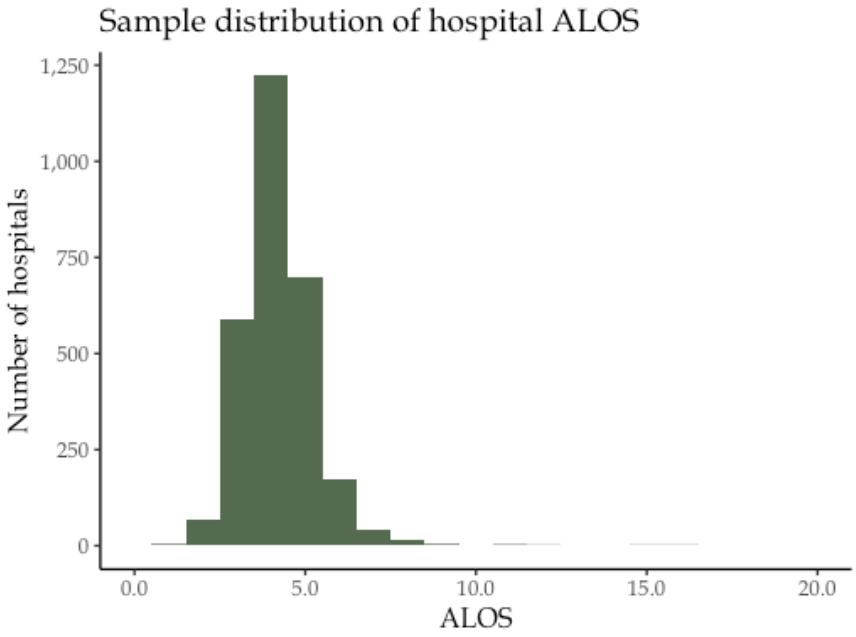


VARIABLES OVERVIEW *distributions*

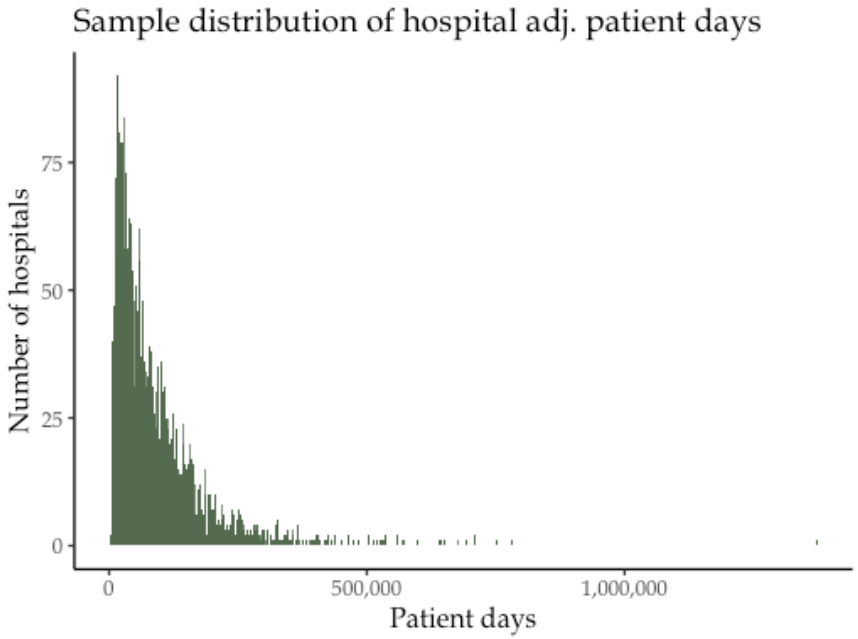


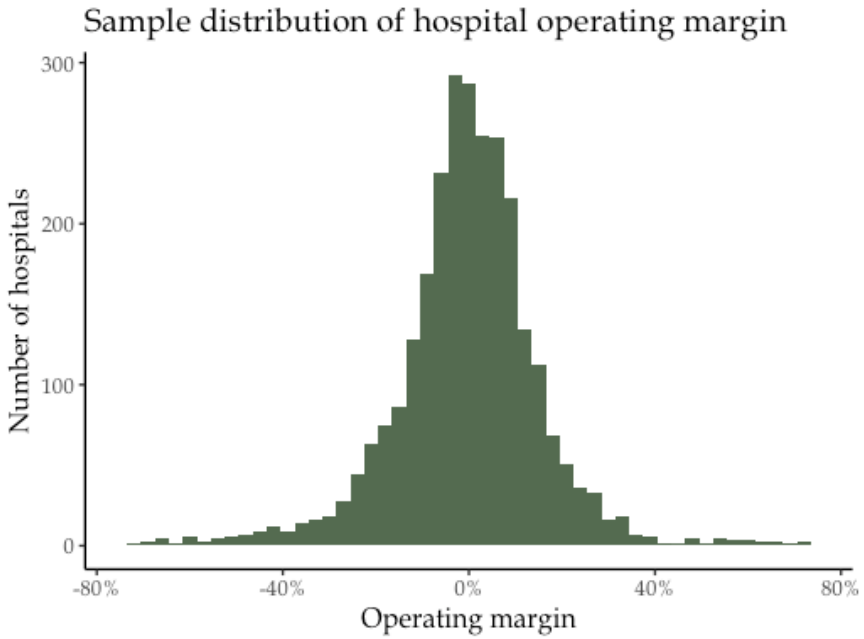


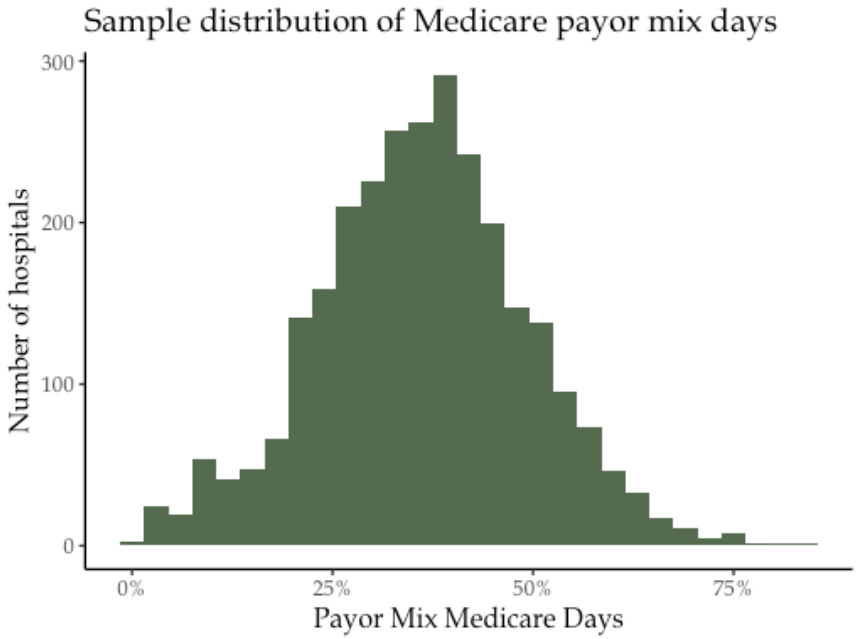




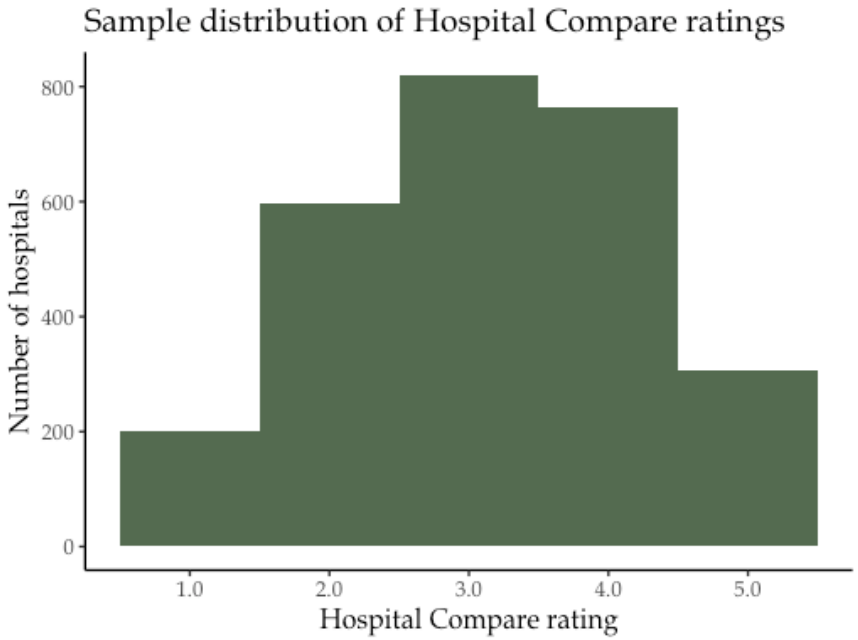
VARIABLES OVERVIEW *distributions*

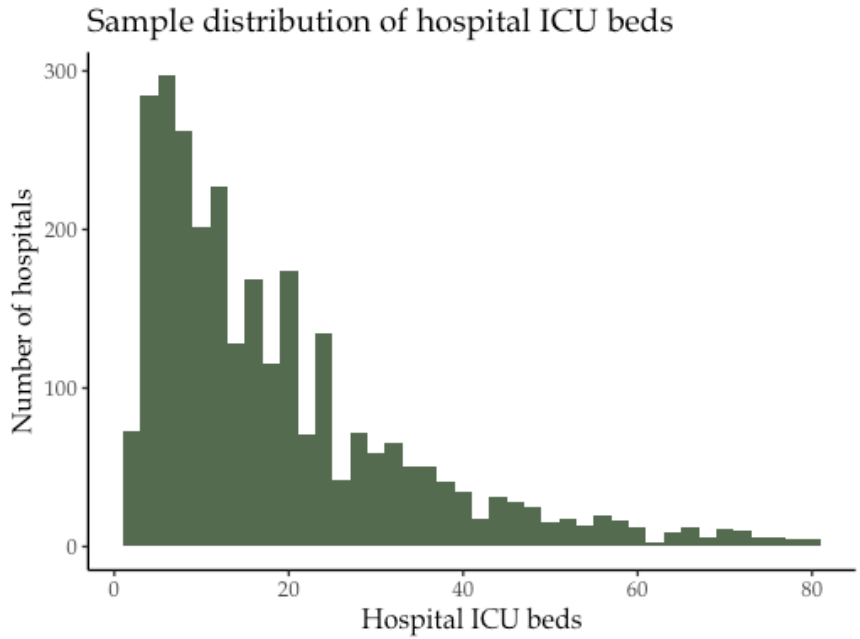


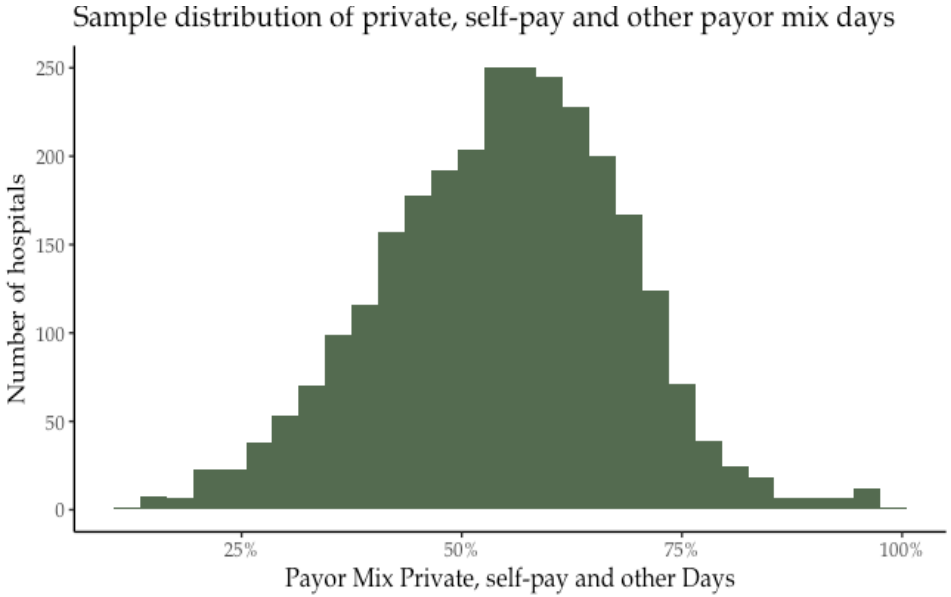




VARIABLES OVERVIEW *distributions*







MULTIPLE LINEAR REGRESSION *data analysis*

Multiple linear regression[1]

Variable	Estimate	Std. error	*t* value	P(>|*t*|)	Significance
Intercept	-14.96	2.68	-5.583	2.60E-08	***
NPR	-6.14E-09	1.20E-09	-5.107	3.51E-07	***
Operating margin	-0.41	0.37	-1.099	0.2719	
EBITDA	1.13E-08	1.80E-09	6.294	3.60E-10	***
Hospital independence	-1.04	0.50	-2.06	0.0395	*
Rural	0.09	0.75	0.123	0.9025	
Payor mix 'Medicare' days	2.01	2.55	0.79	0.4296	
Payor mix 'Private, Self and Other' days	2.68	2.27	1.182	0.2374	
Case mix index	11.91	0.95	12.491	~0	***
Hospital comparison score	-0.52	0.18	-2.881	0.004	**
Discharges	9.38E-04	6.43E-05	14.603	~0	***
Beds	1.65E-02	3.28E-03	5.048	4.77E-07	***
Adjusted patient days	-6.25E-06	6.39E-06	-0.978	0.3283	
ALOS	1.17	0.24	4.868	1.20E-06	***

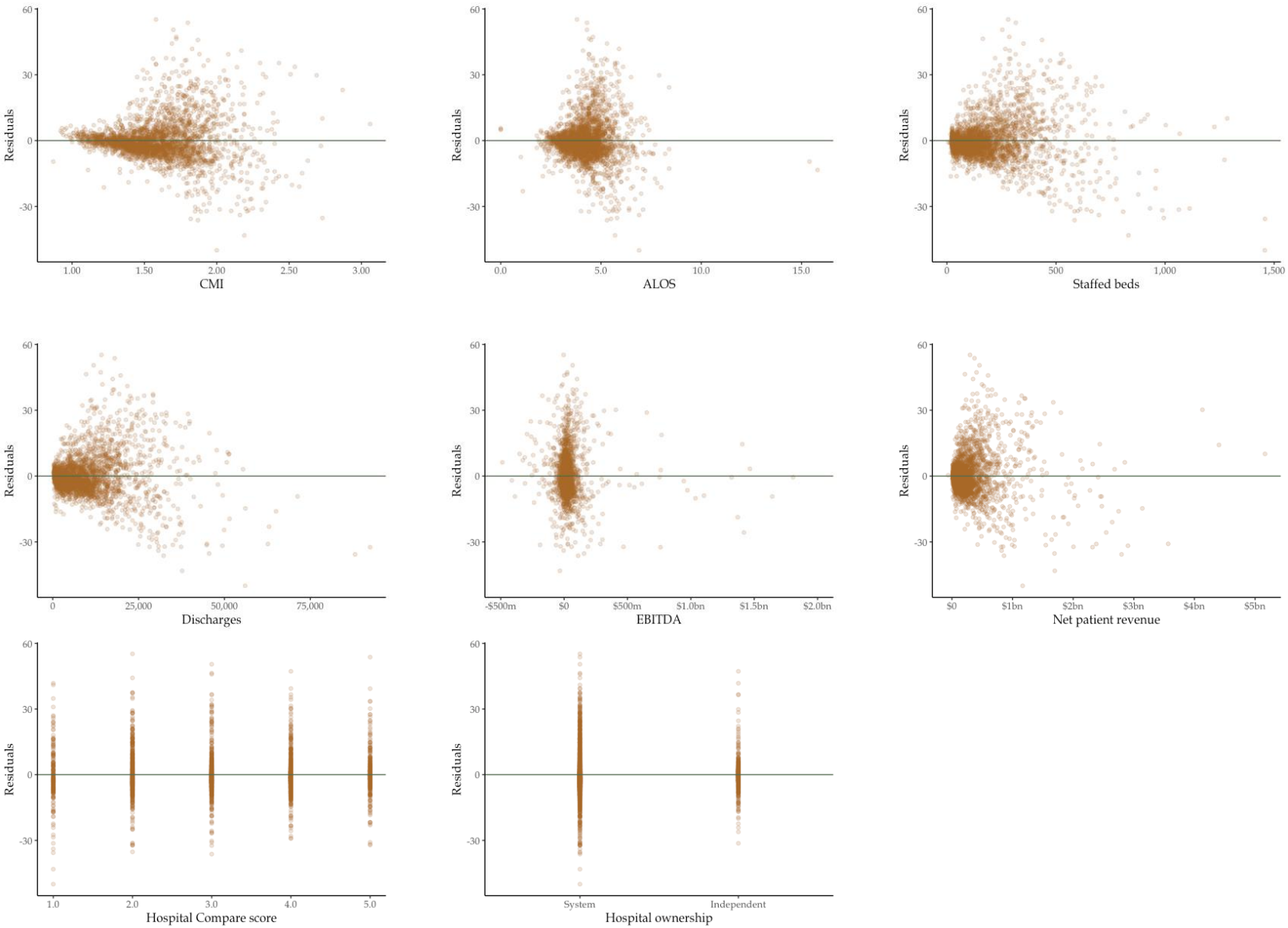
Statistically significant[2]

Linear model

- Coefficients, generated by the model[3] and associated with each variable listed, are given in the 'Estimate' column
  - Those in bold are now considered for the model, but because NPR, EBITDA, Beds and Discharges all capture aspects of size, we can narrow these four to just include Discharges
- Each coefficient's confidence interval is described by the 'Standard error' column
- Strength of the statistical test for significance is described by the 't value' and 'p value' columns

Notes:
1. Notation on E+ used for shortening display of very large and very small numbers; xE+y denotes multiplication of $x*10^{y}$ whereas xE-y denotes multiplication of $x*10^{-y}$
2. * denotes alpha level of 0.05; ** denotes alpha level of 0.01; *** denotes alpha level of 0.001
3. See the appendix for a general description of the form of this type of model

MULTIPLE LINEAR REGRESSION *residual plots*



- Random scatter on the residual plots indicates that the linear model is an appropriate type

MULTIPLE LINEAR REGRESSION *data analysis*

- Variables[1] whose relationship in the model to ICU beds is statistically significant:

Variables	Estimate[2]	Std. error
Intercept	*-11.8900000000*	*1.5460000000*
Hospital independence	-1.4700000000	0.5036000000
CMI	12.4000000000	0.8948000000
Hospital Compare score	-0.7514000000	0.1756000000
Discharges	0.0009488000	0.0000265500
ALOS	1.1920000000	0.2241000000

- Linear model for predicted ICU beds: $\widehat{ICU} = -1.47 \cdot INDEPENDENCE + 12.40 \cdot CMI - 0.75 \cdot HCOMP + 0.000949 \cdot DISCHARGES + 1.19 \cdot ALOS$

- $R^2 = 0.61$ indicates that approximately 61% of the variation in ICU beds can be explained by the linear model above

- For example, let us predict the number of ICU beds at Cleveland Clinic's Indian River Medical Center

Inputs

Case mix index	1.75
Independence	0
Hospital Compare score	3
Number of discharges	13,462
Average length of stay	4.2

Model prediction

Predicted result, *ICU Beds*	24
Actual result	24
Residual (predicted – actual)	*0*

Notes:
1. By assumption, net patient revenue, EBITDA and number of beds and discharges are dependent (perhaps linearly); discharges is the most consistent and accurate figure in the data set; hence, discharges is kept as a variable in the model while the others are excluded
2. Coefficients are shown in the linear model shown lower on the page

COMMENTS *data analysis*

Linear model

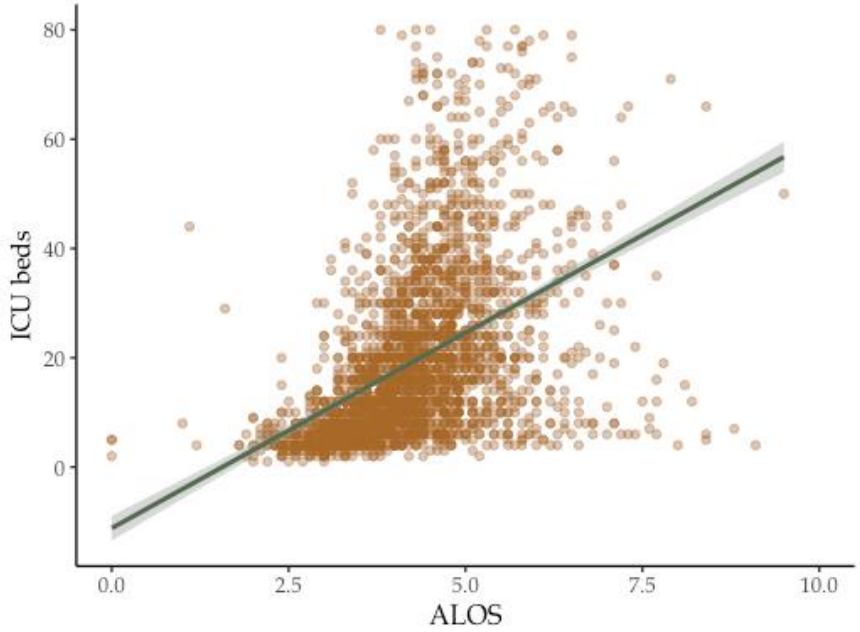
- R^2 = 0.61 reflects a moderate predictive value of the linear model

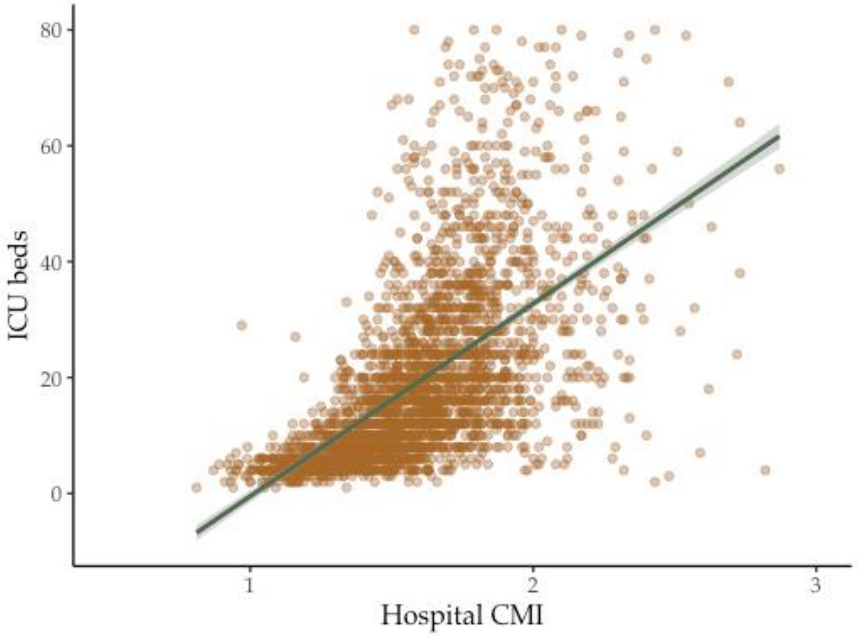
- Coefficients reflect the incremental change in predicted ICU beds associated with an incremental change in an independent variable, *ceteris paribus*:
  - An increase of a hospital's CMI of 1 is associated with an increase of 12 ICU beds
  - A hospital's affiliation[1] with a system is associated with an increase of 1 ½ ICU beds
  - An increase of a hospital's discharges by 1,054 is associated with an increase of 1 ICU bed
  - An increase of a hospital's average length of stay of 9/10 day is associated with an increase of 1 ICU bed

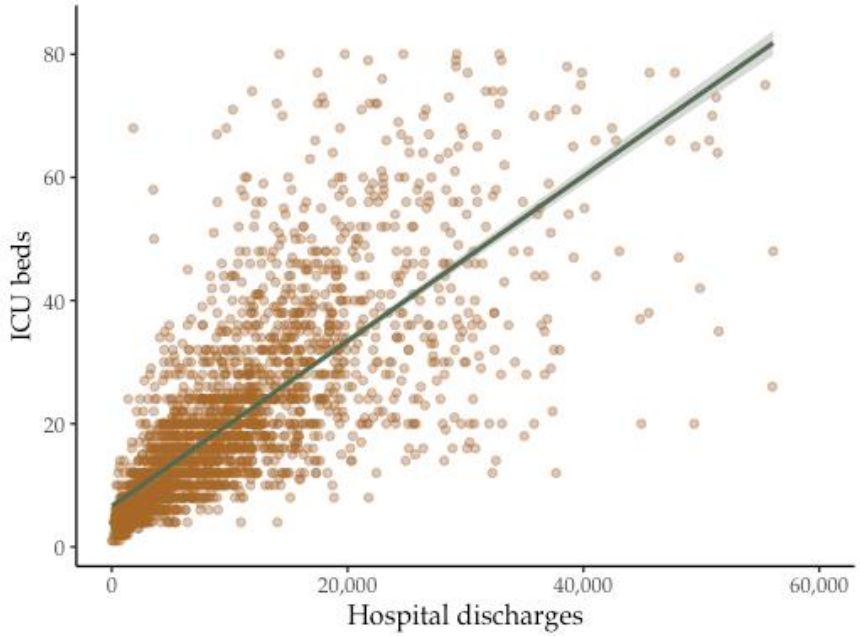
- Random scatter on the residual plots indicate that a linear model is appropriate

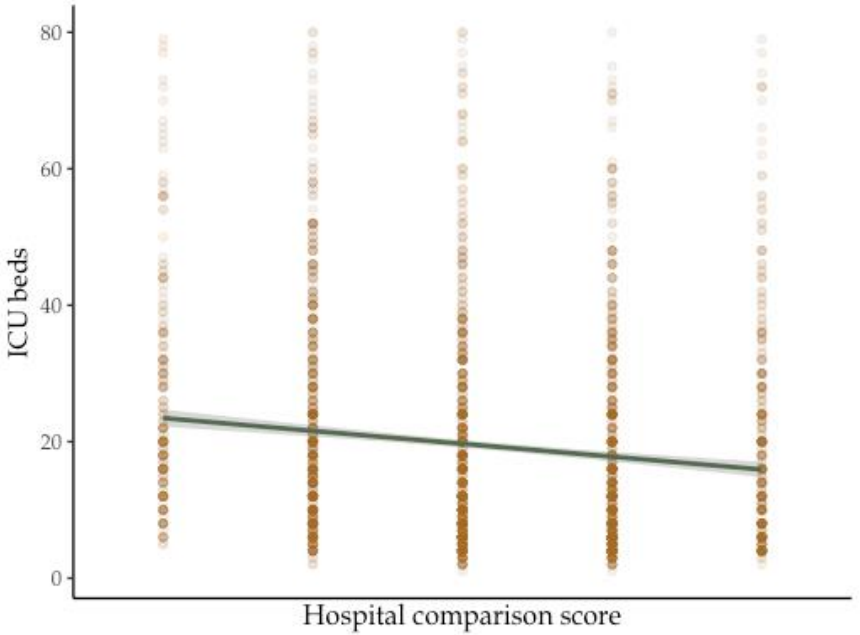
Notes:
1. Defined as having an IDN or IDN parent, per CMS data

ONE-TO-ONE RELATIONSHIPS *continuous variables in the subset*



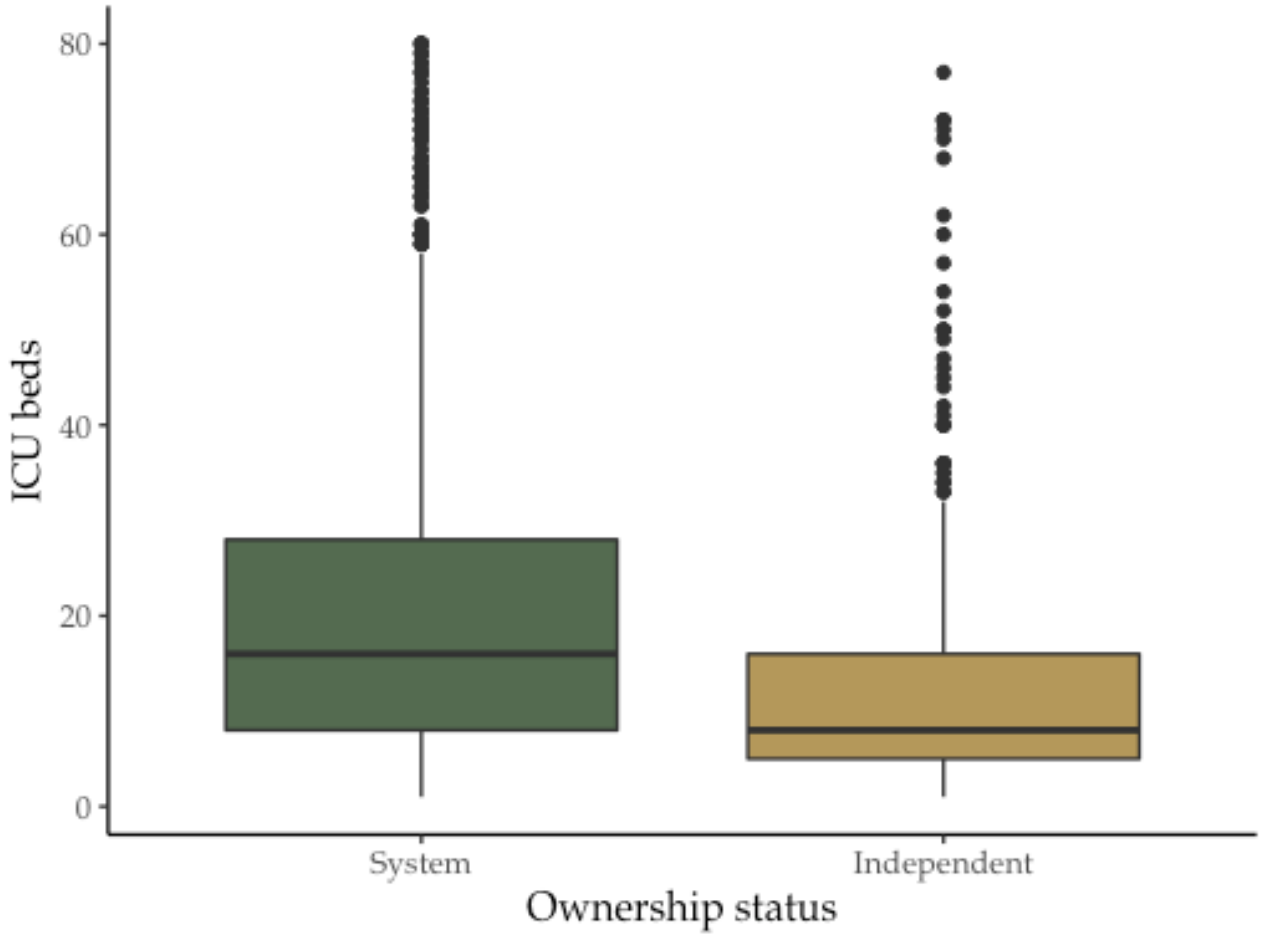






OWNERSHIP STATUS *distribution in the subset*

- Besides the statistically significant relationship, system hospitals generally tend to have more ICU beds, as illustrated below:

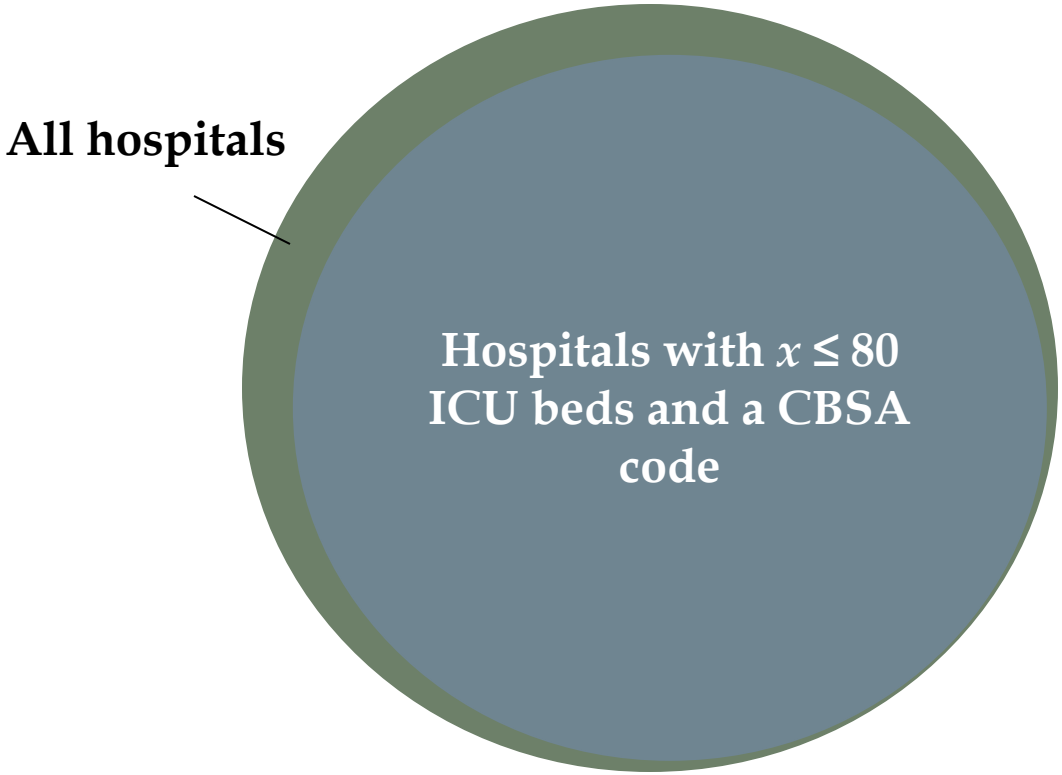


4. Case Mix Index

VARIABLES OVERVIEW *distributions*

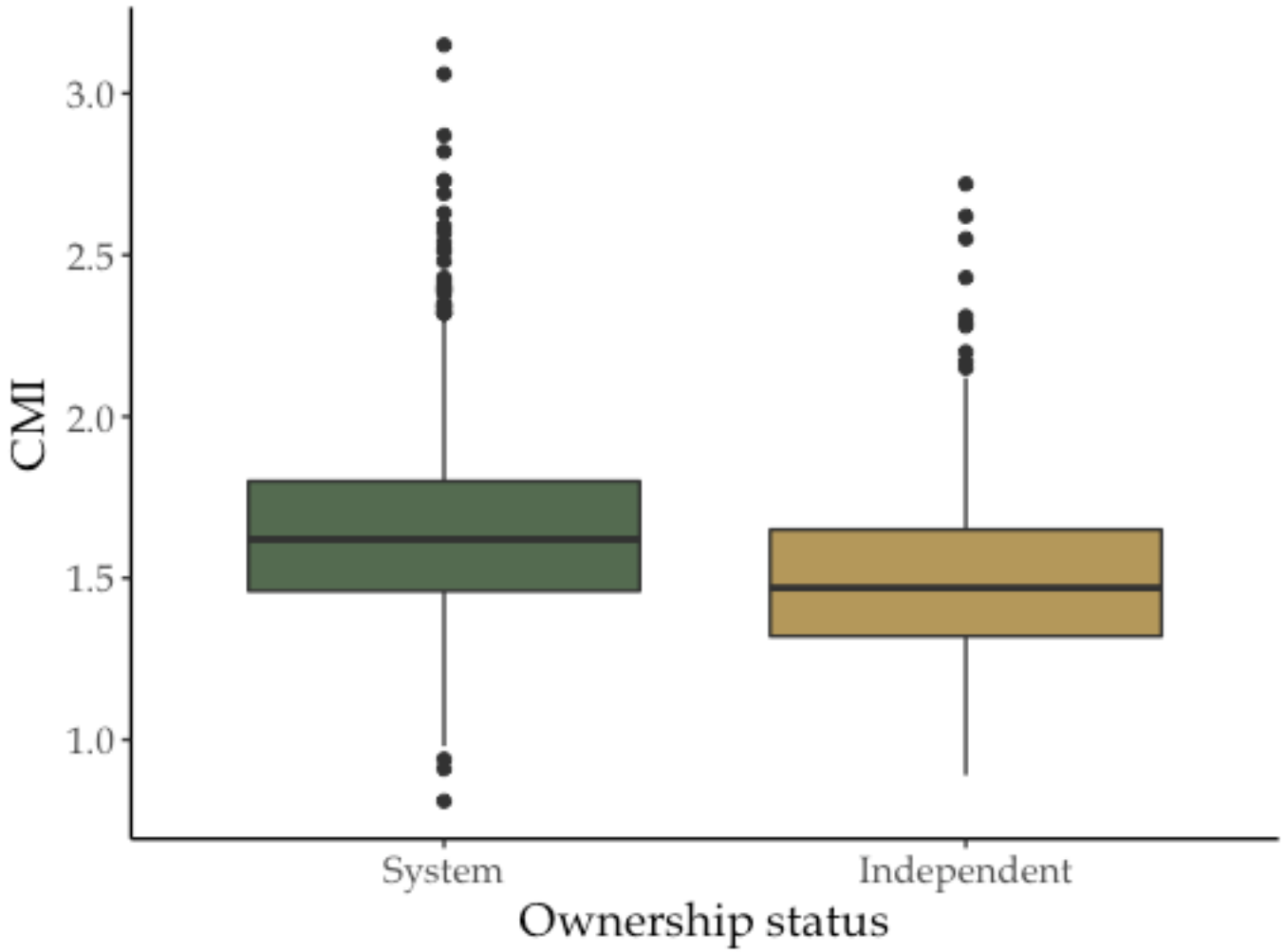
- Here, we consider a subset of the first set: the set of hospitals that have fewer than 80 ICU beds and that have a designated Core-Based Statistical Area (CBSA) code[1]

- CBSA code are assigned by the federal government to areas anchored by an urban center of at least 10,000 people; markets without a CBSA code have low population density



Notes:
1. CMS data

SYSTEM & STANDALONE *distributions*



HOSPITALS IN THE US & CMI *research question*

- Variables[1] whose relationship in the model to CMI is statistically significant:

Variables	Estimate	Std. error
Intercept	1.2550000000	0.0306600000
ICU beds	0.0059400000	0.0003512000
Hospital independence	-0.0553300000	0.0117000000
Payor mix - Medicare days	-0.2144000000	0.0357500000
Hospital Compare score	0.0361200000	0.0037170000
Adjusted Patient Days	0.0000006204	0.0000000589
ALOS	0.0362400000	0.0052730000

- Linear model[2] for predicted CMI:

$$\widehat{CMI} = 1.255 + 0.00594 \cdot ICUBEDS - 0.05533 \cdot INDEPENDENCE - 0.2144 \cdot PAYORMIXMEDDAYS + 0.03612 \cdot + 0.0000006204 \cdot ADJPATIENTDAYS + 0.03624 \cdot ALOS$$

- $R^2 = 0.42$ indicates that approximately 42% of the variation in CMI can be explained by the linear model above

Notes:

1. By assumption, net patient revenue, EBITDA and number of beds and adjusted patient days are dependent (perhaps linearly); for simplicity only adjusted patient days is kept in the model; coefficients are shown in the linear model shown lower on the page
2. The variable Payor mix medicare days is expressed as a percentage; hospital independence is defined here as a Bernoulli random variable, with being associated with a IDN or IDN parent giving a hospital a value of 0, otherwise giving a value of 1

5. Accountable Care Organization and Clinically Integrated Network Participation

SYSTEM & STANDALONE *distributions*

Distribution: proportion of hospitals in ACO		
	System	Independent
ACO	0.53	0.16
No ACO	0.47	0.84

Distribution: proportion of hospitals in CIN		
	System	Independent
CIN	0.28	0.02
No CIN	0.72	0.98

Distribution: proportion of hospitals in ACO or CIN		
	System	Independent
Either	0.81	0.18
Neither	0.19	0.82

HOSPITALS IN THE US & ACO AFFILIATION *research question*

- Variables whose relationship in the model to ACO affiliation is statistically significant:

Variables	Estimate	Std. error
Intercept	*-0.21688*	*0.12370*
Independence	-1.50803	0.12426
Hospital Compare score	0.21254	0.03763

- Logistic model[1] for predicted ACO affiliation:

$$log(\frac{p}{1-p}) = -0.21688 - 1.50803 \cdot INDEPENDENCE + 0.21254 \cdot HCOMP$$

- This logistic model implies that, for a given independent hospital, odds are stronger that it is not in a ACO than for a given system hospital, other characteristics equal

- For example, suppose that a hospital has a Hospital Compare score of 3
  - Then, the model suggests that the probability that it is in an ACO if in a system is 60%
  - And, the model suggests that the probability that it is in an ACO if independent is 25%

- However, the model has relatively weak predictive power[2] and tends to overestimate probabilities for independent hospitals being in ACOs

Notes:
1. Here a logistic regression model is chosen because the dependent variable ACO affiliation is a categorical variable; p is the probability of having a ACO affiliation
2. Given a probability level p, the model tends to have an unremarkable true positive rate and significant false positive rate

HOSPITALS IN THE US & ACO/CIN AFFILIATION *research question*

- Variables whose relationship in the model to ACO affiliation is statistically significant:

Variables	Estimate	Std. error
Intercept	*0.28379*	*0.17741*
Independence	-1.84664	0.12297
Operating margin	0.30008	0.13404
Payor mix Medicare days	0.89474	0.39221
Hospital Compare score	0.14163	0.04112

- Logistic model[1] for predicted ACO/CIN affiliation:

$$log(\frac{p}{1-p}) = 0.28379 - 1.84664 \cdot INDEPENDENCE + 0.14163 \cdot HCOMP + 0.30008 \cdot OPMARGIN + 0.89474 \cdot PAYORMIXMED$$

- This logistic model implies that, for a given independent hospital, odds are stronger that it is not in a ACO/CIN than for a given system hospital, other characteristics equal

- However, given the very low rates of ACO and ICN participation by standalone facilities and high rates of participation for system hospitals shown on the first slide of this section, the model is prone to a high rate of false positives and false negatives

Notes:
1. Here a logistic regression model is chosen because the dependent variable of having either ACO or CIN affiliation is a categorical variable; p is the probability of having either affiliation

APPENDIX

LINEAR REGRESSION *data analysis*

- Four assumptions are required for linear regression analysis
  - Relationship between two variables is linear
  - Expected value of residuals is zero
  - Observations are independent
  - For any value of the independent variable, the dependent variable is normally distributed

- Correlation coefficient and R^2
  - Correlation coefficient (Pearson), or R, is a measure of the strength and direction of the linear relationship between two variables that is defined as the covariance of the variables divided by the product of their standard deviations
    - A value on the range [-1,1]
  - R^2 is the square of the correlation coefficient, otherwise described as the proportion of the variance in the dependent variable that is predictable from the independent variable
    - A value on the range [0,1]

- Form of the model
  - Intercept: a , typically generated by a computer program for a model with multiple variables
  - Coefficients: β_i , i = 1,2,3,… , generated by a computer program
  - Error term: ε
  - Independent variables: x_i , i = 1,2,3,… , selected by the person creating the model
  - Dependent variable: y, the predicted variable the model is designed to predict

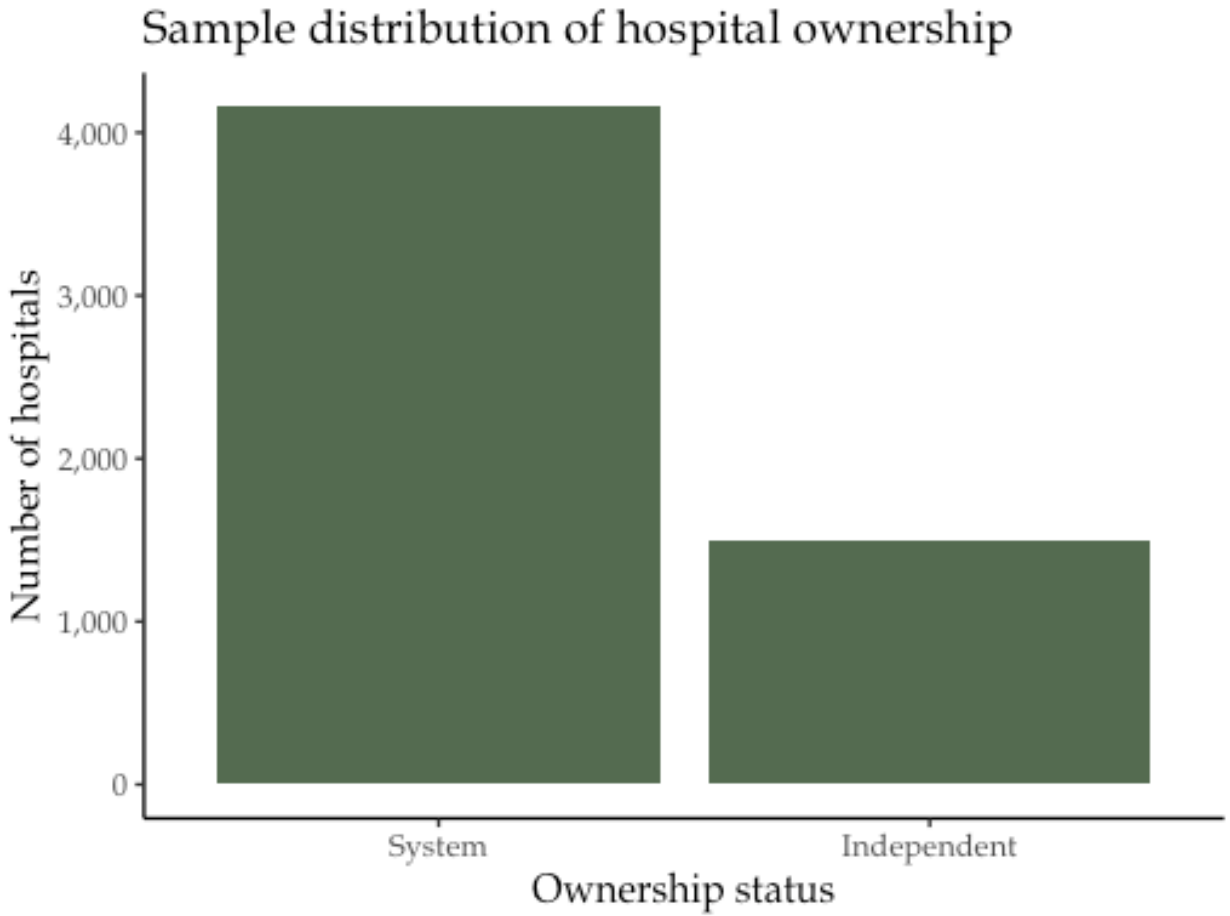
$$\widehat{y} = \alpha + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \ldots + \varepsilon$$

MULTIPLE LINEAR REGRESSION *methods*

- To find out if the relationships between the independent variables and the dependent variable are meaningful, we are interested in their statistical significance
- We examine their coefficients, which describe each variable's relationship
- To test for statistical significance, the program uses a 1-sided *t*-test
- The statistical test generates a p value, which is the probability of obtaining results as extreme as the observed results of a statistical hypothesis test, assuming that the null hypothesis is correct
- The model uses p values to determine statistical significance: if a p value is above 0.05 for a particular hypothesis of a variable, the model excludes the variable

VARIABLES OVERVIEW *data overview*

- Approximately 5.6k facilities in the US, of which:
  - 4.2k are in a system[1]
  - 1.4k are independent



  - The following analysis is based on this set, the set of all critical access and short-term acute-care hospitals in the US instead of the subset

Notes:
1. For the purposes of this analysis, hospital ownership status is a Bernoulli random variable. Hospitals with a IDN or IDN parent are assigned a value of 0 whereas hospitals without either are assigned a value of 1