Supplementary Appendix

A. Calculating Network Measures

We measured two characteristics of the hospital networks examined in this study: median adjusted degree, and PCP relative centrality. Both of these measures were derived from standard, well-accepted network metrics [1,2] with modifications to make them more interpretable for the study of physician networks.

Adjusted Degree

We calculate *degree* by counting the number of the ties a physician has to other colleagues through shared patients. This number is a physician's unadjusted *degree* in the network. For calculating this measure, we do not restrict a physician's ties to her own hospital affiliation but rather consider any ties the physician has within her hospital referral region (HRR). We do not limit physicians' *degree* to shared patients within their affiliated hospital because many physicians share patients with physicians based at other hospitals.

We adjust each physician's *degree* by dividing it by the number of shared Medicare patients seen by the physician in 2006, where the number of shared patients is given by the total number of Medicare patients seen by the doctor whose care was shared with any other treating physician. We perform this adjustment because, for a variety of reasons, there is a wide disparity in the number of Medicare patients seen across physicians. For instance, imagine that two physicians have the same referral patterns and work in identical health care systems, but treat very different numbers of Medicare patients (e.g, one of the physicians may work part time). The physician with fewer shared Medicare patients will always look like she has a lower degree in this analysis despite having identical behavior and similar relationships when compared to the physician with more patients. Due to the fact that the median number of Medicare patients seen in 2006 by the physicians in our data was 120 and because of a previously published study using a similar measure [3], we decided to express the *adjusted degree* as *degree* per 100 Medicare patients.

Lastly, because our analysis was centered on the hospital as a unit of analysis, we summarized the connectivity of physicians at a hospital by taking the median *adjusted degree* of all physicians affiliated with a hospital.

Betweenness Centrality and Relative Centrality

Several measures of centrality exist in the network literature. We chose to rely on the measure of betweenness centrality (the most commonly used measure) due to its applicability in a clinical context and straightforward interpretation. Betweenness centrality is a measure of how likely a member of a network is to be located on the shortest path between any two other members in a network. The equation for betweenness centrality is defined as follows:[2]

$$C_B(n_i) = \sum_{i \neq j \neq k \in N} \frac{g_{jk}(n_i)}{g_{jk}}$$

N = the set of all physicians in a network

n_i = the *i*-th physician

$C_B(n_i)$ = betweenness centrality of physician *i* (n_i)

g_{jk} = the number of shortest paths (also known as "geodesics") which link physicians *j* and *k*

$g_{jk}(n_i)$ = the number of shortest paths which link physicians *j* and *k* that include physician *i*

The expression inside the sum for the equation for $C_B(n_i)$ becomes larger as networks become larger. This is because, in the case of most real-world networks, the number of possible shortest paths in a network grows with the size of a network, so C_B can become arbitrarily large. This makes comparing betweenness centrality values between differently sized networks challenging. A normalization equation has been proposed, but has methodological shortcomings when comparing networks of very different sizes [2].

We chose to bypass this shortcoming of comparing betweenness centrality values across networks by calculating the *relative betweenness centrality*, or *relative centrality* of groups of physicians in a single hospital, i.e. within a given network. The equation for relative centrality is defined as follows (using PCP relative centrality):

$$C_R = \frac{\overline{C}_{H,\,PCP}}{\overline{C}_{H,\,Non-PCP}}$$

$C_R =$ relative PCP centrality for hospital H

$\overline{C}_{H,\,PCP} =$ average betweenness centrality for all PCPs in hospital H

$\overline{C}_{H,\,Non\text{-}PCP} =$ average betweenness centrality for all other physicians in hospital H

The relative centrality, CR, enables us to look at how central one group of physicians is relative to others within a hospital. This value is weakly correlated with the number of physicians of a hospital (r = -0.14 and 0.09 for PCP relative centrality and medical specialist centrality, respectively), versus extremely high correlations using the simple mean betweenness centrality values (r = 0.95 and 0.88 for PCP relative centrality and medical specialist centrality, respectively).

Because some small hospitals in our dataset had mostly physicians who were either PCPs or medical specialists, the denominator for the expression for C_R was 0 and could not be estimated. Likewise, for some hospitals, the denominator was very close to 0 and so C_R became correspondingly large or small. We set these outlying values equal to the 1st and 99th percentiles, accordingly. Sensitivity analyses were performed on our models with various levels of cutoffs for the centrality values, with almost all of the effect of outlying values apparent after trimming the single most extreme centrality value.

B. Statistical Approach

Given the novel application of network measures for understanding physician networks, we did not start our investigation with any theoretical knowledge about the most appropriate functional form of the network measures relative to the outcomes. We decided to keep network measures in their original continuous form because there was no *a priori* science knowledge informing us that a particular categorization would be informative. We were also cautious about categorizing these predictors given the well-documented pitfalls of categorizing continuous predictors [4,5].

We examined the univariate and joint distributions of the network measures used in this study, depicted in **Appendix Figures 1** and **2.** We also measured the centrality of other physician groups, namely medical specialists, surgeons and “other” specialists. Medical specialist relative centrality was significantly correlated with the outcomes studied with similar magnitude, but opposite direction of PCP relative centrality (e.g. high specialist centrality was correlated with higher costs and utilization of care), but those results are not presented for the sake of clarity. Surgeon and other specialist centrality never emerged as statistically significant in any models. There is no substantial correlation between adjusted degree and the relative centrality measures.

We then plotted the two network measures against the 9 outcomes used in this manuscript**.** Overall, no clear nonlinear trends were apparent. With just a few exceptions, our assessment was that all of the network measure versus costs and utilization relationships were well approximated by linear trends. We tested this impression by first entering the network measures as predictors in a multiple linear regression framework. We assessed model fit and the assumption of linearity of each of the predictors with thorough diagnostic checks, including partial residual plots. In the small number of cases where the plots might have suggested a different functional form, we also fit models using logarithmic, quadratic and cubic transformations of the network measures, but did not find that any consistently provided a better fit than the original measure. Therefore, we favor leaving the network measures as linear predictors, enabling simpler, more interpretable, and, most likely, more robust results.

C. Details of Multiple Regression Models

1.

In building our multiple regression models, we began with an interest in using the detailed hospital-level performance data available from the Dartmouth Atlas of Health Care as a set of outcome measures [6]. We then performed a focused review of recent literature on hospital costs and quality and defined a set of control variables that we felt encompassed a wide array of likely confounders that could explain variation in hospital costs and health care utilization independently of network measures [7-10]. Details on the data sources and which databases contained which variables are explained in the Methods section of the paper and in **Appendix Table 1**, below. One variable we eliminated from our pre-specified group of covariates was a "technology index" which gave each hospital a score based on the level of advanced medical technology (e.g., robotic surgery) available at a hospital. We eliminated this variable because it has a high prevalence of missing data (~20%) and its inclusion in the model, with or without multiple imputation for missing data, made no substantive change in the results of our analyses.

We fit a linear, log-normal multiple regression model with the following equation:

$$\begin{aligned}\ln(\mathrm{HospitalOutcome}) &= \alpha + \beta_0 \mathrm{AdjustedDegree} + \beta_1 \mathrm{CPRelativeCentrality} \\ &+ \boldsymbol{\beta}^T \mathrm{HospitalCovariates} + \varepsilon\end{aligned}$$

Separate models were fit for each 9 hospital outcomes. The hospital outcomes, network measures, and hospital covariates are summarized in **Appendix Table 1**.

To account for the fact that smaller hospitals had fewer patients in the last 2 years of life for which to calculate the hospital outcomes used by the Dartmouth Atlas group, we weighted our regression models by the average number of annual deaths at each hospital during the period over which the outcomes data were gathered, 2001-2005. In addition, to account for the possibility that the variance of an observation varies with the corresponding mean or with values of the predictors, we used robust heteroscedastic-consistent standard error estimation procedures [11,12].

D. Sensitivity Tests and Model Diagnostics

We thoroughly evaluated the 9 models using a wide array of regression diagnostics. We found evidence of mild heteroscedasticity in the distribution of the residuals for several of the models, so we used robust heteroscedastic standard error estimation [11,12].

Due to the number of covariates used in our regression models, we were concerned that the counterfactuals underlying the reported effects presented in **Fig. 3** in the manuscript may be extrapolations not supported by actual observations in our data. In particular, we present the counterfactuals of the expected change in hospital outcomes given a change of 1 standard deviation in a network characteristic for the average non-profit, urban, non-teaching, hospital in our dataset. To insure that we were not reporting effects involving counterfactuals with no or minimal support from the data, we calculated the multi-dimensional Gower's distance of these counterfactuals from our entire dataset [13,14]. We found that for every effect we present in **Figure 3**, increasing or decreasing any of the 2 network measures by 1 standard deviation, is within 0.1 Gower distance away from 89 or more other hospitals in the dataset [15]. In summary, with sophisticated diagnostics proposed by King and Zeng, we found that the effects we report in **Figure 3** are based on conservative, well-supported interpolations within our dataset.

We conducted a wide range of sensitivity tests on our set of 9 models. First, we tested the effects of the categorization of hospitals into 2 or 3 categories of urban/rural location (either urban/non-urban or urban/rural/isolated) and teaching status (either teaching vs. non-teaching or major/minor/non-teaching). There was no substantive change to the results with either set of categorizations. We also added covariates to assess the influence of patient population, looking at the effect of adding mean patient age and percentage of black or white patients to our models. In addition, we examined the potential effect of Medicare Advantage penetration in regional markets on our models by calculating the percentage of patients enrolled in Medicare Advantage for at least one month in 2006 in each HRR and included this covariate in our model. The addition of these variables made no substantive difference to any of the results in the models. The addition of these variables also made no substantive difference to any of the results in the models.

Sensitivity analyses on outlying data were also performed because we were concerned about their possible effect on our results. In particular, some models had a few observations whose residual values were quite large or small. We performed sensitivity analyses assessing the impact of these outliers by excluding a few extreme outliers, and also excluding the outer 1st-10^{th} percentile of residual outliers in our data from the models. We found that these changes had either a minimal effect or increased the estimates and precision of our results. We opted for the more conservative presentation of the results obtained by not excluding any outlying data. As described in part B above, we also performed sensitivity analyses on the effect of trimming the centrality ratios at various levels of cutoffs (1st-10^{th} outer percentiles) for the centrality values, with almost all of the effect of outlying values apparent after trimming the single most extreme centrality value. We decided to set the most extreme centrality ratios to the 1^{st} and 99^{th} percentile values.

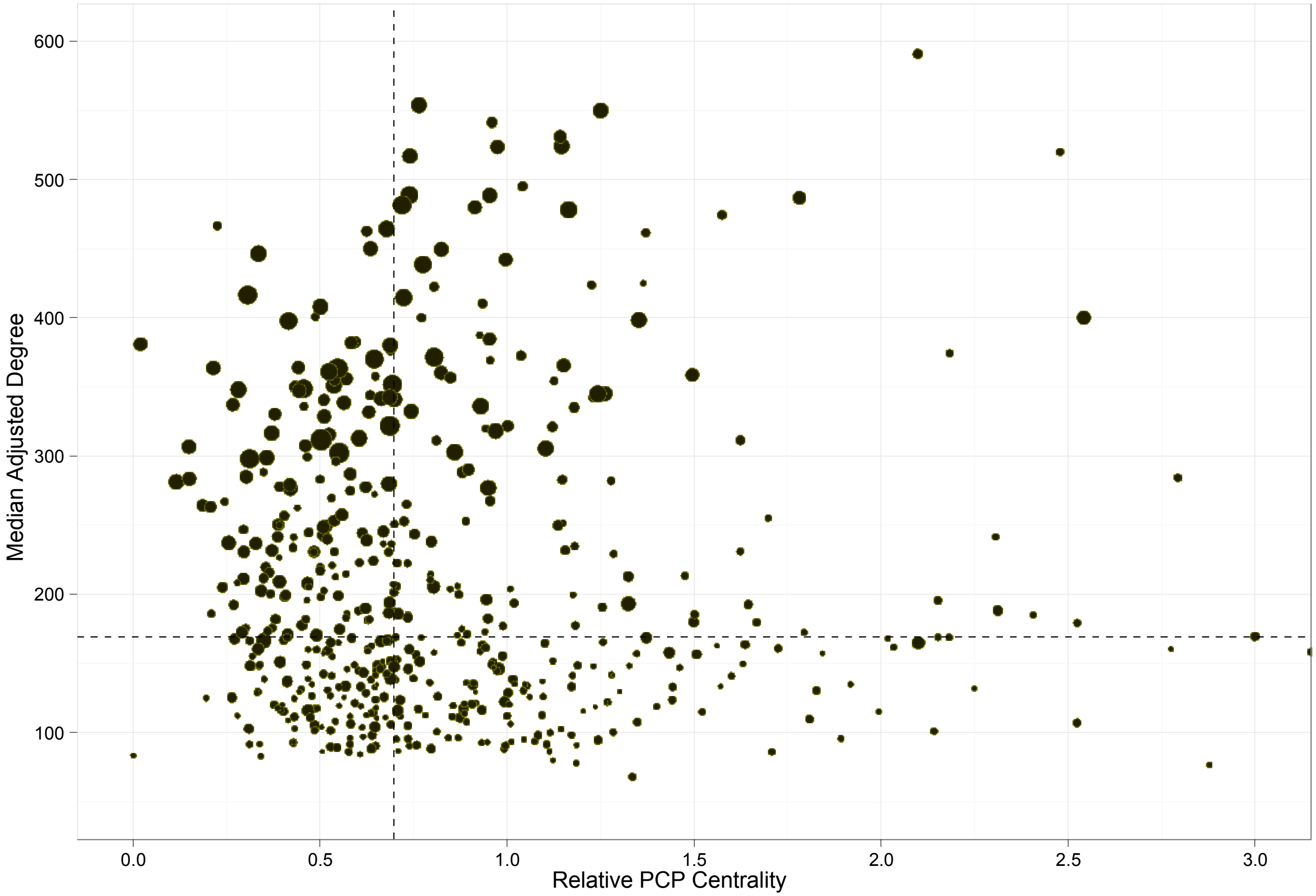
Lastly, we simulated the effect of a hypothetical unobserved confounder on the results of our models. We created a series of simulated binary unobserved confounders for all models in the manuscript with a range of associations with both the network measures and outcomes of interest. The associations with both outcome and network predictor ranged from a log odds ratio of -3 to a log odds ratio of 3, a range far greater than one might expected for a predictor for which we and others have no knowledge of to date. We simulated an unobserved confounder with an average prevalence of 50%. We found that in general, most models were not sensitive to an unobserved confounder without a very strong association with both the outcome and confounder. Exceptions included median adjusted degree vs. medical/surgical days and PCP centrality vs. total costs and total PCP visits. All of these associations had confidence intervals closest to including 0 among the significant results presented in **Fig. 3**. In these models, a binary unobserved confounder with a log odds ratio association with a standard deviation change in the network predictor or outcome of +/- 0.5 to 1.0 (exact values depending on the model) could take away the statistical significance of the result.

Overall, we found that the majority of our statistically significant results held up to unobserved confounding, unless the unobserved confounder was strongly associated with both the predictor and the outcome (with far stronger associations than seen with any of the measured confounders we included in our analyses). The results that were more sensitive to unobserved confounding (discussed above) were the weaker associations presented in the results. It is possible an unobserved confounder exists that could be similar to the confounder we simulated, but in most cases it is improbable that we and others are completely unaware of a factor that has such a strong association to both cost/utilization outcomes and network properties.

Appendix References

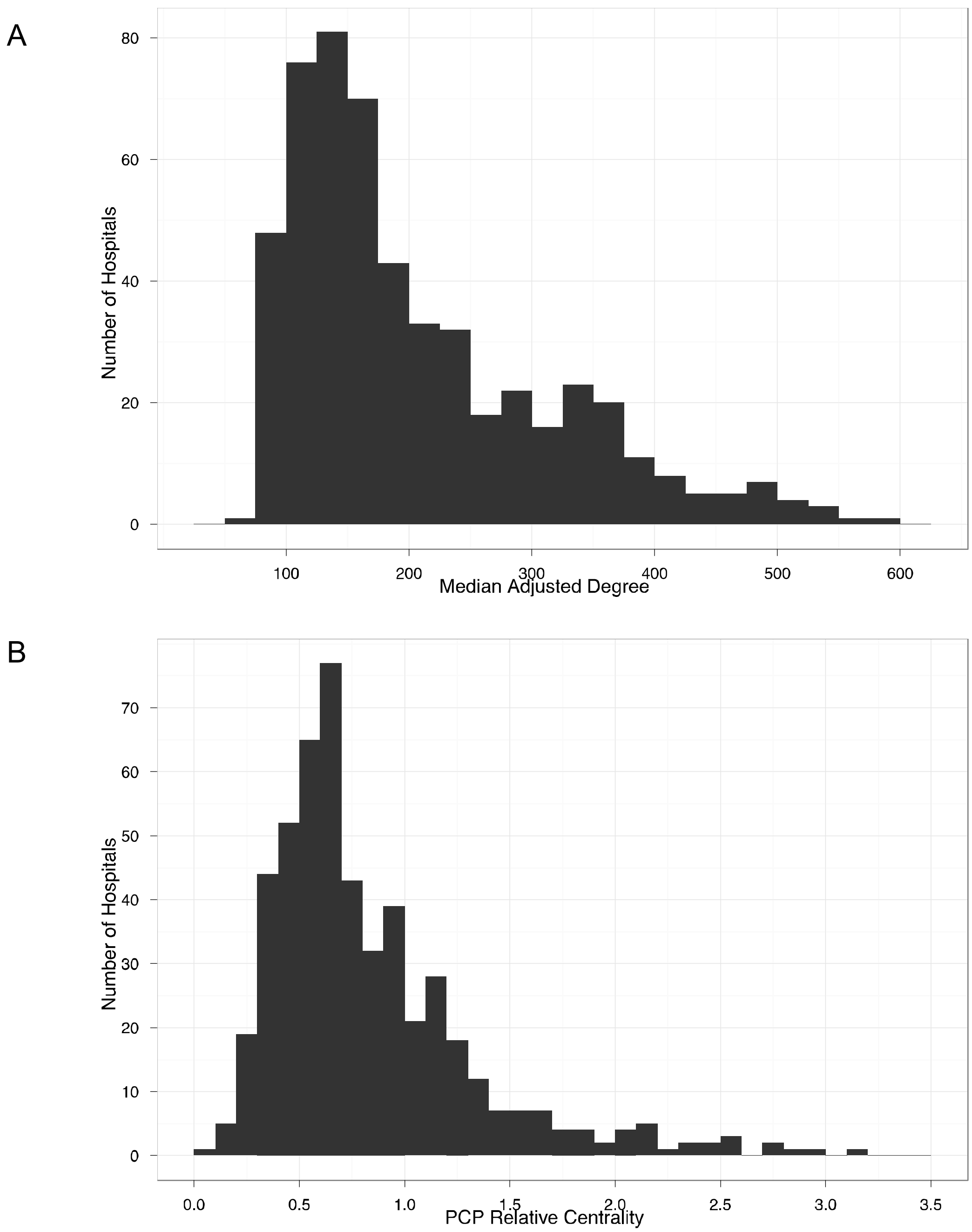
1. Newman M. The structure and function of complex networks. SIAM Reviews 2003.
2. Wasserman S, Faust K. Social Network Analysis: Methods and Applications. Cambridge, MA: Cambridge University Press; 1994.
3. Pham HH, O'Malley AS, Bach PB, Saiontz-Martinez C, Schrag D. Primary care physicians' links to other physicians through Medicare patients: the scope of care coordination. Annals of internal medicine 2009;150:236-42.
4. Royston P, Altman DG, Sauerbrei W. Dichotomizing continuous predictors in multiple regression: a bad idea. Stat Med 2006;25:127-41.
5. Problems Caused by Categorizing Continuous Variables. 2010. (Accessed October 4, 2010, 2010, at http://biostat.mc.vanderbilt.edu/wiki/Main/CatContinuous.)
6. Wennberg JE, Fisher E, Goodman DC, Skinner J. Tracking The Care of Patients with Severe Chronic Illness: The Dartmouth Atlas of Health Care, 2008. Lebanon, NH: The Dartmouth Institute for Health Policy & Clinical Practice, Center for Health Policy Research,; 2008.
7. Jha AK, Orav EJ, Dobson A, Book RA, Epstein AM. Measuring efficiency: the association of hospital costs and quality of care. Health affairs (Project Hope) 2009;28:897-906.
8. Landon BE, Normand S-LT, Lessler A, et al. Quality of care for the treatment of acute medical conditions in US hospitals. Archives of internal medicine 2006;166:2511-7.
9. Barnato AE, Chang CC, Farrell MH, Lave JR, Roberts MS, Angus DC. Is survival better at hospitals with higher "end-of-life" treatment intensity? Medical care 2010;48:125-32.
10. Yasaitis L, Fisher ES, Skinner JS, Chandra A. Hospital quality and intensity of spending: is there an association? Health affairs (Project Hope) 2009;28:w566-72.
11. Zeileis A. Econometric Computing with HC and HAC Covariance Matrix Estimators. Journal of Statistical Software 2004;11:1-17.
12. White H. A heteroskedasticity-consistent covariance matrix and a direct test for heteroskedasticity. Econometrica 1980;48:817-38.
13. King G, Zeng L. The Dangers of Extreme Counterfactuals. Political Analysis 2006;14:131-59.
14. King G, Zeng L. When Can History be Our Guide? The Pitfalls of Counterfactual Inference. International Studies Quarterly 2007;51:183-210.
15. Stoll H, King G, Zeng L. WhatIf: Software for Evaluating Counterfactuals. In. R package version 1.5-5 ed; 2009.
16. Bynum JP, Bernal-Delgado E, Gottlieb D, Fisher E. Assigning ambulatory patients and their physicians to hospitals: a method for obtaining population-based provider performance measurements. Health services research 2007;42:45-62.

Appendix Figure 1: Hospital Network Characteristics vs. Total Medicare Costs



In this figure, each point represents a hospital. The size of each point corresponds to each hospital's total Medicare spending. The x-axis corresponds to the relative primary care provider (PCP) centrality in that hospital, and the y-axis corresponds to the median adjusted degree of physicians in that hospital. Dashed lines are drawn at the median values of relative PCP centrality and median adjusted degree to guide the eye. In this figure, 18 hospitals (of 521 total with non-missing values) with high PCP centralities > 3.2 fall outside the range of the plot. The concepts of adjusted degree and PCP centrality are presented in **Fig. 1C** and **Fig. 2**. The Pearson correlation coefficient of median adjusted degree versus PCP relative centrality is -0.04 (p = 0.35).

Appendix Figure 2: Univariate Distributions of Hospital Network Measures



This figure shows histograms of the distribution of network measures across the 528 hospitals in the dataset used for this study. In **App. Fig. 1B**, one outlying centrality value (which equals approximately 17) falls outside the range of the plot.

Appendix Table 1: List of Hospital Outcomes, Network Measures and Hospital Covariates

Hospital Outcomes (patients in last 2 years of life)	Network Measures	Hospital Covariates
Data Source: Dartmouth Atlas of Health Care	**Data Source:** Medicare Claims	**Data Sources:** American Medical Association (AMA), American Hospital Association (AHA), or derived from Medicare Claims
Total Costs Imaging Costs Test/Laboratory Costs Total Hospital Days General Medical/Surgical Days ICU Days Total Physician Visits PCP Visits Medical Specialist Visits	Median Adjusted Degree PCP Relative Centrality	Number of physicians (Medicare, see ref [16]) Number of hospital beds (AHA) Number of RN FTE's per 1000 inpatient days (AHA) Percentage of Medicare admissions (AHA) Percentage of Medicaid admissions (AHA) Urban or non-Urban (AHA) Major, Minor or Non-Teaching (AHA) Non-profit, For-profit or Public (AHA) Mean patient volume of physicians (Medicare) Percentage of PCPs (AMA)

Appendix Table 2: Complete Regression Coefficient Tables

Outcome	Total Costs			Imaging Costs			Test/Lab Costs		
	Adj. R^2	F-statistic	p-value	Adj. R^2	F-statistic	p-value	Adj. R^2	F-statistic	p-value
	0.559	48	<0.001	0.563	48.9	<0.001	0.489	36.5	<0.001
	Beta	**SE**	**p-value**	**Beta**	**SE**	**p-value**	**Beta**	**SE**	**p-value**
Intercept	10.90	0.016	0.000	6.71	0.022	0.000	6.19	0.027	0.000
Median Adjusted Degree	0.164	0.016	0.000	0.214	0.021	0.000	0.195	0.026	0.000
PCP Relative Centrality	-0.062	0.011	0.000	-0.097	0.015	0.000	-0.138	0.018	0.000
Number of physicians	-0.003	0.013	0.816	0.067	0.017	0.000	0.068	0.021	0.001
Number of beds	-0.016	0.011	0.126	-0.021	0.014	0.146	-0.042	0.017	0.014
Number of RN FTE's per 1000 inpatient days	-0.052	0.009	0.000	-0.048	0.013	0.000	-0.087	0.016	0.000
Mean patient volume of physicians	0.012	0.012	0.327	0.110	0.017	0.000	0.125	0.020	0.000
Percentage of Medicare admissions	0.023	0.011	0.042	0.027	0.015	0.084	0.025	0.018	0.171
Percentage of Medicaid admissions	0.032	0.011	0.003	-0.011	0.015	0.441	0.024	0.017	0.162
Percentage of PCPs	-0.047	0.013	0.000	-0.071	0.017	0.000	-0.084	0.021	0.000
Minor teaching hospital (Reference: None)	-0.023	0.019	0.234	-0.116	0.027	0.000	-0.088	0.032	0.005
Major teaching hospital (Reference: None)	0.024	0.030	0.417	-0.197	0.041	0.000	-0.178	0.049	0.000
Public hospital (Reference: Non-profit)	-0.030	0.023	0.189	-0.020	0.031	0.514	-0.082	0.037	0.029
For profit hospital (Reference: Non-profit)	0.028	0.025	0.275	0.070	0.035	0.044	0.088	0.042	0.036
Non-urban hospital (Reference: Urban)	-0.045	0.024	0.063	-0.136	0.033	0.000	-0.145	0.040	0.000

Outcome	Hospital Days			General Medical or Surgical Days			ICU Days		
	Adj. R^2	F-statistic	p-value	Adj. R^2	F-statistic	p-value	Adj. R^2	F-statistic	p-value
	0.472	34.3	<0.001	0.299	16.7	<0.001	0.283	15.5	<0.001
	Beta	**SE**	**p-value**	**Beta**	**SE**	**p-value**	**Beta**	**SE**	**p-value**
Intercept	3.10	0.018	0.000	2.19	0.027	0.000	1.24	0.040	0.000
Median Adjusted Degree	0.160	0.017	0.000	0.082	0.026	0.002	0.315	0.038	0.000
PCP Relative Centrality	-0.043	0.012	0.000	-0.035	0.019	0.060	-0.058	0.027	0.031
Number of physicians	-0.061	0.014	0.000	-0.045	0.021	0.031	-0.105	0.031	0.001
Number of beds	0.059	0.012	0.000	0.056	0.018	0.001	0.086	0.026	0.001
Number of RN FTE's per 1000 inpatient days	-0.115	0.010	0.000	-0.120	0.016	0.000	-0.079	0.023	0.001
Mean patient volume of physicians	0.037	0.013	0.006	-0.020	0.020	0.322	0.155	0.029	0.000
Percentage of Medicare admissions	0.052	0.012	0.000	0.070	0.019	0.000	0.003	0.027	0.917
Percentage of Medicaid admissions	0.035	0.012	0.003	0.045	0.018	0.011	0.026	0.026	0.323
Percentage of PCPs	-0.045	0.014	0.002	-0.071	0.021	0.001	0.020	0.031	0.519
Minor teaching hospital (Reference: None)	-0.026	0.021	0.223	0.008	0.032	0.808	-0.058	0.047	0.219
Major teaching hospital (Reference: None)	0.011	0.033	0.743	0.116	0.049	0.019	-0.166	0.072	0.022
Public hospital (Reference: Non-profit)	-0.029	0.025	0.252	-0.090	0.038	0.018	0.074	0.055	0.185
For profit hospital (Reference: Non-profit)	-0.031	0.028	0.268	-0.179	0.042	0.000	0.241	0.062	0.000
Non-urban hospital (Reference: Urban)	0.085	0.027	0.002	0.182	0.040	0.000	-0.174	0.059	0.003

Outcome	Total Physician Visits			PCP Visits			Medical Specialist Visits		
	Adj. R^2	F-statistic	p-value	Adj. R^2	F-statistic	p-value	Adj. R^2	F-statistic	p-value
	0.477	34.9	<0.001	0.301	17	<0.001	0.557	47.7	<0.001
	Beta	**SE**	**p-value**	**Beta**	**SE**	**p-value**	**Beta**	**SE**	**p-value**
Intercept	4.20	0.019	0.000	2.78	0.022	0.000	2.61	0.034	0.000
Median Adjusted Degree	0.213	0.018	0.000	0.161	0.021	0.000	0.338	0.032	0.000
PCP Relative Centrality	-0.090	0.013	0.000	-0.064	0.015	0.000	-0.159	0.023	0.000
Number of physicians	-0.050	0.015	0.001	-0.051	0.017	0.003	-0.090	0.026	0.001
Number of beds	0.028	0.012	0.025	0.022	0.014	0.121	0.060	0.022	0.006
Number of RN FTE's per 1000 inpatient days	-0.093	0.011	0.000	-0.099	0.013	0.000	-0.132	0.020	0.000
Mean patient volume of physicians	0.108	0.014	0.000	0.042	0.017	0.011	0.231	0.025	0.000
Percentage of Medicare admissions	0.031	0.013	0.018	0.079	0.015	0.000	-0.021	0.023	0.369
Percentage of Medicaid admissions	0.014	0.012	0.251	0.038	0.014	0.008	0.016	0.022	0.456
Percentage of PCPs	-0.030	0.015	0.046	0.066	0.017	0.000	-0.141	0.026	0.000
Minor teaching hospital (Reference: None)	-0.062	0.023	0.006	-0.022	0.026	0.399	-0.092	0.040	0.022
Major teaching hospital (Reference: None)	-0.104	0.035	0.003	-0.021	0.040	0.601	-0.120	0.061	0.051
Public hospital (Reference: Non-profit)	-0.083	0.027	0.002	-0.108	0.031	0.001	-0.129	0.047	0.007
For profit hospital (Reference: Non-profit)	0.044	0.030	0.140	-0.009	0.035	0.791	0.141	0.053	0.008
Non-urban hospital (Reference: Urban)	-0.037	0.028	0.188	0.027	0.033	0.407	-0.231	0.050	0.000