The Influence of Physician Referral Networks on Submitted Charges for Cataract Surgery

Team members: Song Wang, Daniel Ricci, Ruosi Guo

1. Abstract

Medical services are extremely expensive across the U.S. and make up a large proportion of household spending. In this paper, we use multiple public data sets released by the Centers for Medicare and Medicaid Services, including payment, physician referral, physician info, etc. to model the charge submitted by physicians for a specific service -- cataract surgery. More specifically, we want to see whether the roles physicians play in the regional network have effects on the charge submitted.

If a physician plays a more central role in network, indicating he is prestigious, he/she may tend to charge more for a service than less prestigious peers. We built two types of models to explore this question. One model uses domain knowledge to construct a linear regression model. None of the three regional network measures selected were significant in the final model and poorly explained the observed variance (R^2=0.05). In further exploration we introduced some tree based methods to build a prediction model. It confirms the finding we have in linear regression -- node centrality is not significant factor in predicting the submit amount.

2. Introduction

Previous research as well as popular media outlets have documented the seemingly outrageous prices that are being charged for medical services. News articles describing itemized hospital charges of $25 aspirin tablets, $100 dollar diapers, and $18 blood sugar tests have popped up from a variety of news sources recently.[1,2] Each of these charges impressively exceed the actual costs associated with the products and services provided. Research has observed charges for these services in the neighborhood of 2.5 to 3.7 times higher than their respective costs.[3,4] These costs are commonly defined by the amount the government run health insurance program, Medicare, will pay for a given service.

U.S. citizens who are age 65 and older are eligible for government provided health insurance coverage. This insurance coverage, known as Medicare, is managed by the Centers for Medicare and Medicaid Services (CMS) and consists of four major parts: Part A (hospital insurance), Part B (medical insurance), Part C (Medicare

Advantage Plans), and Part D (prescription drug coverage).[5] For the purposes of this study, we will focus exclusively on Part B - medical insurance - which covers services provided by doctors, outpatient care, medical supplies, and preventative services. Cataract surgery falls under this category.

The purpose of the study is to explore the average charges submitted to Medicare for one such service, cataract surgery, and examine whether it is related to measures of physician prestige, characterized by centrality measures within regional referral networks, physician characteristics while controlling for the influence of regional cost and demographic differences. In short, do more prestigious physicians charge more for their services?

2.1 Widening Gap Between Charges and Payments for Medical Services

The gap between charged amount and payment received is thought to be a consequence of the evolution of health care system with the divergence beginning in the mid-1980s. At this time two concurrent systems of payment existed (1) government determined rates and (2) privately negotiated rates between insurance companies and providers.[6] Initially, Medicare reimbursement was profitable, but over time, in an effort to lower medical costs, Medicare began lowering its reimbursement for services. To accommodate this financial pressure providers began charging more to private payers. The evolution of this gap since 1992 is illustrated in Figure 1.[7]

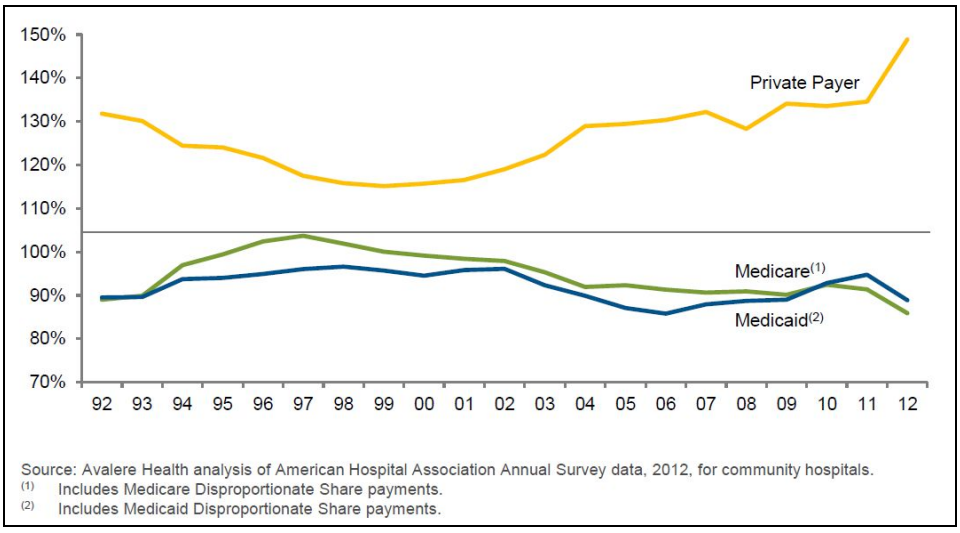

Figure 1: Hospital Payment-to-cost Ratios for Private Payers, Medicare, and Medicaid, 1992 - 2012

2.2 Medicare Payment Structure in the U.S. Healthcare System

To better understand submitted charges and Medicare payments an explanation of Medicare payment structure is warranted. There are three main parties involved in the provision and procurement of Part B services and the structure is outlined in Figure 2. First a provider, in this case an ophthalmologist, performs a service (cataract surgery) and submits a bill to Medicare, this is known as the submitted charge amount. Medicare then determines the total amount of money the provider will receive, this is known as the Medicare allowed amount for a given service. Each service has a base price, the physician fee schedule, which is then regionally adjusted for cost of living differences which determines the total allowed amount.[8] The adjustment equation used by Medicare for the year 2012 is shown in Figure 3. Finally, after the allowed amount is determined Medicare pays 80% of this cost and the patient is responsible for the remaining 20%.[9]

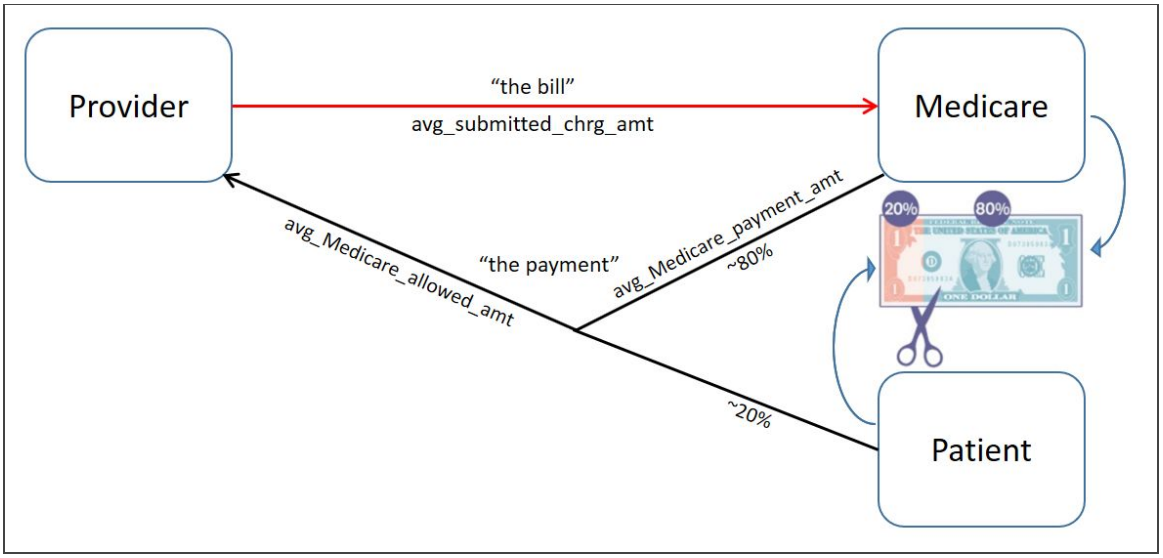

Figure 2: Medicare Payment Structure. Providers submit charges to Medicare for services they perform and this "bill" is known as the submitted charge amount. Medicare determines how much it will pay by calculating a base cost and adjusting it for regional cost variability. This total is known as the Medicare allowed amount. For the most part, 80% of this allowed amount is paid by Medicare and the remaining 20% is provided by the patient.

$$Payment_{H,L} = [(GPCI_{PW,L} \times RVU_{PW,H}) + (GPCI_{PE,L} \times RVU_{PE,H}) + (GPCI_{MP,L} \times RVU_{MP,H})] \times CF$$

Physician Work | Practice Expense | Malpractice Insurance

GPCI	Component Index	Measures Geographic Differences in:
Physician Work	Single Component	Physician wages
Practice Expense	*Employee Wage*	Wages of clinical and administrative office staff
	Purchased Services	Cost of contracted services (e.g., accounting, legal, advertising, consulting, landscaping)
	Office Rent	Physician cost to rent office space
	Equipment, Supplies, and Other	Practice expenses for inputs such as chemicals and rubber, telephone use and postage
Malpractice	Single Component	Cost of professional liability insurance

Figure 3: Medicare Payment Regional Variability

$Payment_{H,L}$ - Medicare allowed amount for a given service, H, in Medicare locality, L, after regional adjustment.
GPCI - Geographic Practice Cost Indices for a given Medicare locality, L; Measure geographic differences in input prices for the three parts: physician work (PW), practice expense (PE), and malpractice insurance (MP).
RVU - Relative Value Units; Relative importance and value of each of the three parts: PW, PE, and MP.
CF - Conversion Factor; Indicates the dollar value assigned to a RVU in a given locality, L.
Each of the three GCPIs are comprised of a component index containing one to four unique indicators each.

2.3 The Focus on Submitted Amount

As described above, any regional variability in Medicare allowed amount, and subsequent Medicare payment amount would be almost entirely due to the regional adjustment variables found in Figure 3. In addition, these prices are only set by Medicare, and a provider's network prestige would not appreciably influence Medicare's payment determination for services. However, a quantity that could be influenced by a physician's prestige would be the submitted charge amount for a specific service. For example, ophthalmologists who occupy more prestigious network positions may be inclined to submit higher charges for their services than their less prestigious peers.

2.4 Determination of Submitted Amount

In order to investigate whether a physician's regional network prestige would influence the amount submitted for services, it is important to first understand how submitted charges are set. The prices that hospitals and providers

charge for given services is contained in one centralized computer file known as a charge description master or chargemaster. The goal of the chargemaster is to cover the cost of resources and operational costs while maximizing revenue.[6] Guidance is available on how to set, maintain, and update a chargemaster, however, previous research demonstrates that the chargemaster determination process is inconsistent among providers.[10,11] Additionally, chargemasters are not published nor accessible to the public.[7] Utilizing the information available on the chargemaster construction[6,11] and knowledge gained from an interview with an expert on the subject.[12] We were able to create a conceptual model to guide our analysis (Figure 4).

2.5 Theoretical Model for Provider Submitted Charge Amount

We assume that providers determining their submitted amounts will consider regional cost adjustments much like the adjustments performed by Medicare. Provider cost is intended to account for the provider's wage from the equation in Figure 3 above. Utilities, facility, etc. and Raw materials encompasses the practice expense index from above. Malpractice insurance information was not available in the data available nor was a suitable proxy identified, so it was excluded from the model. Finally, the notion of Service fee was added to reflect the price of the service provided. This domain is thought to be influenced by the provider's regional network prestige. The logic is that prestigious physicians provide a better service and therefore may decide to charge more because of it.

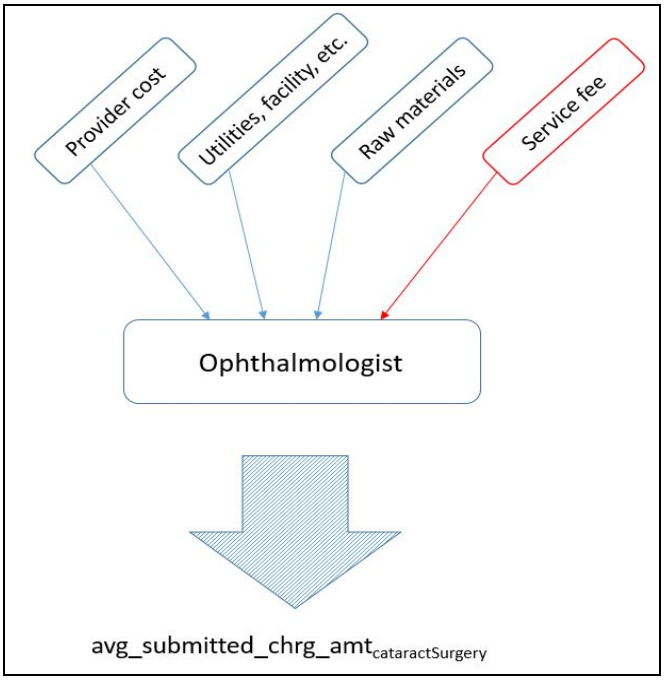

Figure 4: Theoretical Model

2.6 Objective

The further objective is : 1) to use data to verify that this model is valid in term of modeling fitting; 2) test the hypothesis that the referral network among the physicians is an important factor in determining the submit amount -- prestigious physicians, characterized by with a relative higher "importance" in the network will tend to charge more.

3. Statistical Modeling

Four publicly available datasets utilized in this study. Variables of interest were extracted from each dataset, processed, and analyzed to address our goals.

3.1 Data Description

The first dataset used in this study was the Medicare Fee-For-Service Provider Utilization & Payment Data (2012).[13] This data set is prepared CMS and includes data on services and procedures provided to Medicare beneficiaries by physicians and other suppliers during the 2012 calendar year. Information on utilization, payment and submitted charges for these services is organized by National Provider Identifier(NPI), healthcare common procedure coding system (HCPCS) code, and place of service. The physician and other supplier public use file includes data for providers that had a valid NPI and submitted Medicare Part B

non-institutional claims. Records for a service only appear if rendered for at least ten unique beneficiaries. From these data ophthalmologists performing cataract surgery were selected. Cataract surgery was thought to be a good candidate because it is a commonly performed procedure among the Medicare eligible population and since it is a non-life-threatening condition and patients have the ability to choose their surgeon, it would be a good candidate to uncover effects of regional prestige.

We also used the Physician Shared Patient Patterns 30-day (2012) dataset.[14] These data are also provided by CMS and contains data on referrals that occur within 30 days between providers for the 2012 calendar year. The data were used to construct referral networks for physicians. The *referrals* are more appropriately characterized as *shared patient patterns*. If a single patient is seen by a physician and then, within 30 days, is seen by different physician it appears as a link in these data. Much like the previous dataset, these links only appear in instances of greater than ten occurrences during the year.

The third dataset used in this study was the Physician Compare Downloadable Database.[15] This database contains information on active Medicare providers and includes name, medical credentials, medical practice information, and Medicare participation information.

The last dataset used in this study was from the American Community Survey (ACS) 2012 1-Year Estimates file.[16] The ACS is a supplement to the decennial Census that is administered to approximately 3.5 million randomly selected addresses each year. These data are publically available and cover demographic, social, economic, and housing subjects.

There are a lot messy things going with the data sets, like missing data, mismatches, etc. After a long-time spent cleaning the data, we reduced our 22449 observations for cataract surgery, to 7703 cases in the final data set.

3.2 Exploratory Analysis

The nation-wide referral network is too complex to visualize as it contains 1 million nodes and 70 million edges. Instead we examine the regional network for each hospital referral region. To exhibit what such networks look like, we plot the HRR that includes Madison, WI (Figure 5).

Regional network features, like vertice_count, edge_density etc, are some summary/ aggregation for this HRR network. Ophthalmologist's node centrality will be calculated also based on HRR networks, like closeness_centrality, page_rank etc.

Figure 5: Network for one HRR including Madison, WI. The red large-sized dots are ophthalmologists, while the black small-sized dots are all other physicians. There are 2621 physicians, 54 ophthalmologists. There are very limited connections among ophthalmologists themselves.

Each physician within the larger HRR network also possess an ego network identifying the 30-day shared patient patterns for that physician in the 2012 calendar year (Figure 6).

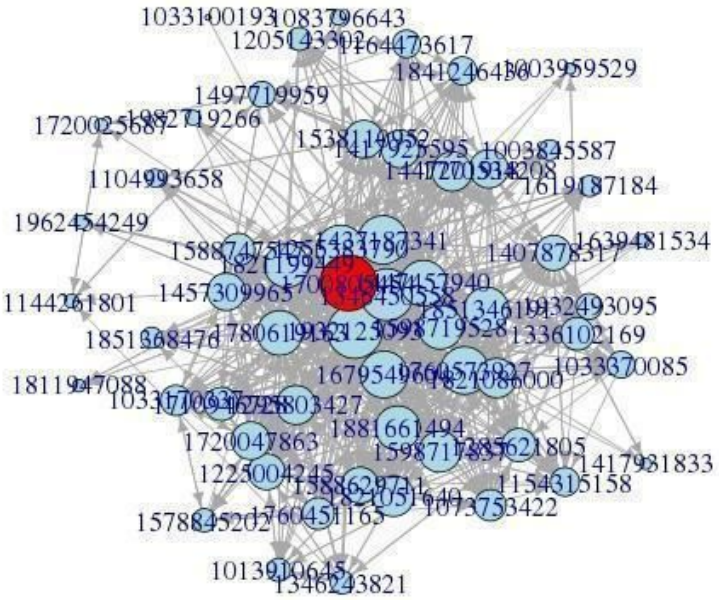

Figure 6: Ego network of one ophthalmologist in Madison, WI. The red dot is ophthalmologist and blue dots are the other physicians referred patients to /from this ophthalmologist. The size reflect the scaled version of degree in the ego-network.

3.3 Linear Model

For simplicity and interpretability, we think use linear regression as the tool to explore the relationship between submitted_amount and the various predictors. We built multiple linear regression model with different features in and out. The results are shown in Table 1. Our models involve two level of features, node-level and HRR-level, and node centrality features are of interest to our project goal.

From Table 1, gender is not significant, while yrs_exp has significant negative effect -- which contradicts to our intuition. This could be because older physicians take less Medicare patients and also are co-surgeons. Also MedicalGrossRent and nonWhite are significant model.

Also, many of regional network features are significant in the modeling procedure. When regional network features as control are in the model, the node

centrality (used several alternatives) are not significant. This will be contradictory to our hypothesis since we claim that control the region level differences, node with high centrality score should be significant.

Final model will be the second one based on our analysis. But from the R^2 score =0.05, we cannot fully convinced that this is a good model. This may be due three reasons: one is there are quite some outliers in the response, which big error; two is that the surrogates for quantities we need the model are not accurate enough; three the model is not valid and we miss something important. We cannot say at this moment.

Table 1: Regression Results

	Dependent Variable:				
	submitted_amount				
	Model 1 (se)	Model 2 (se)	Model 3 (se)	Model 4 (se)	Model 5 (se)
Physician Characteristics					
phys_male	-39.48 (29.99)	-28.12 (29.82)	-27.46 (29.90)	-28.70 (29.83)	-29.32 (29.87)
yrs_exp	-19.68 (11.20)*	-23.25 (11.14)**	-23.05 (11.16)**	-23.15 (11.14)**	-23.34 (11.14)**
Service Volume					
number_eligible	-4.62 (11.28)	-12.14 (11.27)	-12.12 (11.27)	-11.67 (11.28)	-12.13 (11.27)
number_procedures	-138.62 (11.07)***	-134.17 (11.04)***	-133.79 (11.11)***	-133.80 (11.05)***	-135.25 (11.15)***
Regional Cost Controls					
medHHincome	-10.54 (19.37)	-23.78 (19.40)	-23.91 (19.40)	-23.73 (19.40)	-23.71 (19.40)
medGrossRent	97.75 (19.20)***	77.69 (19.71)***	77.69 (19.71)***	78.27 (19.72)***	77.75 (19.71)***
Regional Demographic Control					
non_white	62.36 (12.33)***	46.75 (12.36)***	46.53 (12.38)***	46.92 (12.36)***	46.99 (12.36)***
Regional Network Controls					
vertex_count		109.75 (14.98)***	109.94 (15.00)***	105.13 (15.93)***	109.96 (14.99)***
edge_density		- 232.01 (36.76)***	-233.12 (36.94)***	-226.07 (37.41)***	-230.92 (36.79)***
degree_mean		86.45 (18.48)***	90.28 (22.39)***	65.68 (30.53)**	84.13 (18.79)***
degree_SD		222.46 (36.78)***	222.24 (36.79)***	220.80 (36.83)***	222.70 (36.78)***
Regional Network Prestige					
page_rank			-5.14 (16.94)		
closeness_centrality				19.92 (23.30)	
betweenness_centrality					8.06 (11.76)
Intercept	2093.76 (27.20)***	2084.33 (27.03)***	2083.78 (27.10)***	2084.81 (27.04)***	2085.32 (27.07)***
Observations	7703	7703	7703	7703	7703
Adj R^2	0.039	0.052	0.052	0.052	0.052
F statistic (df)	45.35***	39.23***	35.97***	36.02***	36.00***

	(df= 7;7695)	(df=11;7691)	(df= 12;7690)	(df= 12;7690)	(df= 12;7690)

* p<0.05; **p<0.01; ***p<0.001. All predictors are standardized to facilitate comparison.

3.4 Alternative non-linear approach

After applying our theory based model we decided to consider this problem from a pure prediction standpoint to determine the charge amount a physician will submit for cataract surgery. From this perspective, we will be less constrained by the domain knowledge and low quality of data available. Instead, the non-linear method may allow us to more accurately predict the submitted charge amount and thus provide more insight in the submitted charge amount decision-making process.

Here we try decision tree as our alternative.The decision tree gives better fitting compared with regression approach. The random Forest is a ensemble version of tree methods, which gives a reasonably good fit (see Table 2). We also display a tree given by decision tree algorithm in Figure 6.

Table 2. Results from multiple approaches

	linear regression	decision tree	random Forest
R^2	0.05	0.10	0.32
$\sqrt{MSE}$	2073.6	929.1	809.2

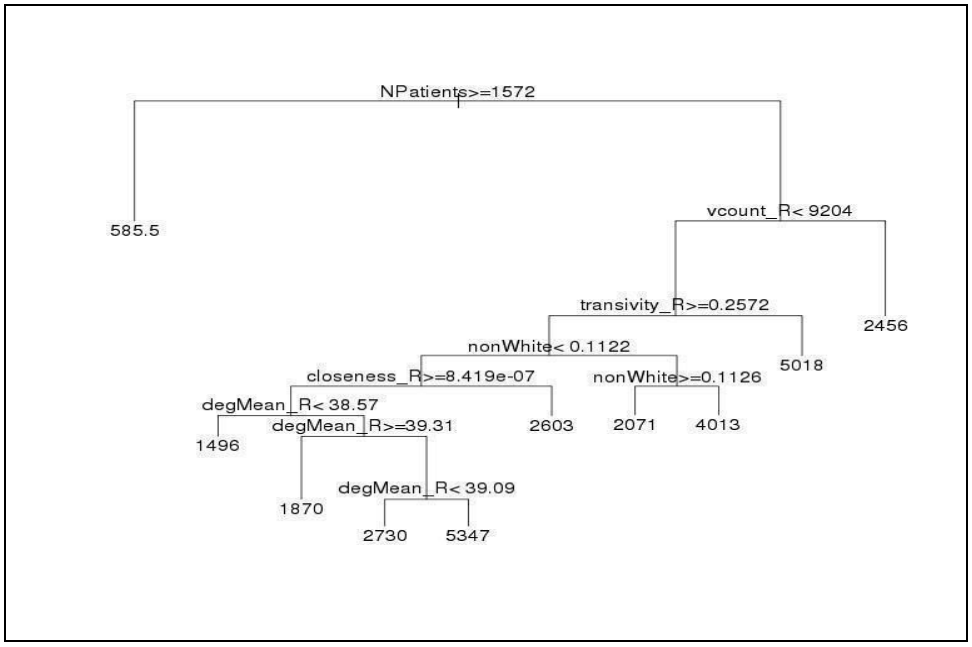

Figure 7: Results from Decision Tree Analysis. In the tree variables work as split can important variable. Here several graph features shows up as split point (those end with '_R' indicates regional version)confirms the significant result from regression.

4. Discussion and Concerns

We conducted a statistical analysis to address the goal -- whether node centrality of physicians in the referral network will be an important factor in determining their submitted charge amount for cataract surgery. However, there are some limitations of methods which needs to be mentioned.

1. Network feature calculations -- node features are calculated within each Hospital Referral Region. As we know, many network features like centrality of each node will characterize some property relative to their networks. And it may be not meaningful to compare two such measures directly. One potential solution would be to include the different HRRs as factor variables (more than 200 levels), or using some form of regional control variables instead. This argument may not be convincing enough especially if the regional control no sufficient.

2. As mentioned before, from the fitting perspective, the linear regression is extremely poor fitting result. R^2 is just around 0.05 for the regression. This may be due to multiple reasons as mentioned. This lack of fit is likely due to the quality of the variables we were able to obtain from these publically available datasets. Using proxy variables such as median household income to approximate physician hourly wages likely resulted in the low R^2 values observed.

3. In the late analysis, we also find there are some outliers both in response and predictor variables. Like the number of procedures performed during the year. The largest values from this variable suggest that ~30 ophthalmologists were performing <u>at least</u> 30 procedures every day of the year.

5. Conclusions

In this project, we have explored about the interesting problem of modeling submitted charge amount for cataract surgery by ophthalmologists. We have gathered public data sets from multiple sources, extracted useful features out of them. We have used two approaches to tackle this problem. The first approach was theory-based (or domain knowledge based), we proposed a theory based on domain knowledge; and we use the best surrogate variables for the quantities in our model. Using linear regression, we were be able to see how variables of interest will affect

the response. The other approach was a decision tree based prediction, examining which variables of interest have importance in predicting the response.

The regression model concludes that after we control regional income, facilities cost as well as regional-level network features , we did not get significant result for the node centrality score. Therefore, we do not believe that an ophthalmologist's location within a hospital referral region has a substantial influence on the charge amount submitted for cataract surgery.

The prediction model given by decision tree or random forest also show limited predictive power of node centrality score compared to other control variables. However, the tree results do suggest that volume does play a role in discounting the charge submitted for cataract surgery.

To make our findings more valid and sound, we may need to do more investigation, especially to address the potential limitations or concerns brought up during the modeling process.

References

1. Why an aspirin taken in a hospital can cost upwards of $25. Healthcare Finance News. http://www.healthcarefinancenews.com/blog/why-aspirin-taken-hospital-can-cost-upwards-25. Accessed December 18, 2015.

2. Bitter Pill: Why Medical Bills Are Killing Us | TIME.com - 2013,2,26,MedicalCostsDemandAndGreed.pdf. http://www.uta.edu/faculty/story/2311/Misc/2013,2,26,MedicalCostsDemandAndGreed.pdf. Accessed December 18, 2015.

3. Anderson GF. From “Soak The Rich” To “Soak The Poor”: Recent Trends In Hospital Pricing. *Health Aff*. 2007;26(3):780-789. doi:10.1377/hlthaff.26.3.780.

4. Bai G, Anderson GF. Extreme Markup: The Fifty US Hospitals With The Highest Charge-To-Cost Ratios. *Health Affairs*. 2015;34(6):922-928. doi:10.1377/hlthaff.2014.1414.

5. What’s Medicare? | Medicare.gov. https://www.medicare.gov/sign-up-change-plans/decide-how-to-get-medicare/whats-medicare/what-is-medicare.html. Accessed December 16, 2015.

6. Tompkins CP, Altman SH, Eilat E. The Precarious Pricing System For Hospital Services. *Health Aff*. 2006;25(1):45-56. doi:10.1377/hlthaff.25.1.45.

7. Avalere Health. *Trendwatch Chartbook 2014: Trends Affecting Hospitals and Health Systems*. American Hospital Association; 2014.

8. MaCurdy T, Shafrin J, DeLeire T, et al. Geographic Adjustment of Medicare Payments to Physicians: Evaluation of IOM Recommendations. *Burlingame, CA: Acumen, LLC*. 2012. http://cms.hhs.gov/Medicare/Medicare-Fee-for-Service-Payment/PhysicianFeeSched/Downloads/Geographic_Adjustment_of_Medicare_Physician_Payments_July2012.pdf. Accessed December 8, 2015.

9. Part B costs | Medicare.gov. https://www.medicare.gov/your-medicare-costs/part-b-costs/part-b-costs.html. Accessed December 17, 2015.

10. Educational CyberPlayGround, Inc.®. How does a hospital decide how much you pay and what to bill you for? *cyberplayground*. https://cyberplayground.wordpress.com/2013/05/23/do-not-pay-the-hospital-bill/. Accessed December 18, 2015.

11. American Health Information Management Association. Care and Maintenance of Chargemasters (Updated). *Journal of AHIMA*. March 2010. http://library.ahima.org/xpedio/groups/public/documents/ahima/bok1_047258.hcsp?dDocName=bok1_047258. Accessed December 18, 2015.

12. Kharbat, Mohammad, PharmD, BCPS. "Hosptial Submitted Charge Amount." Personal interview. 13 Nov. 2015.

13. Centers for Medicare and Medicaid Services. Medicare Physician and Other Supplier Public Use File Methodology. September 2015. https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/Medicare-Provider-Charge-Data/Downloads/Medicare-Physician-and-Other-Supplier-PUF-Methodology.pdf.

14. Centers for Medicare and Medicaid Services. Physician Shared Patient Patterns Technical Requirements. 2015.

https://downloads.cms.gov/foia/physician_shared_patient_patterns_technical_requirements.pdf.

15. Centers for Medicare and Medicaid Services. Physician Compare Downloadable Database Dictionary. 2014. https://data.medicare.gov/views/bg9k-emty/files/h7l0QrAdAmn9kPZ_wabQDjD9Ph-j0rWwR_ewuVlTJd8?filename=Physician_Compare_Downloadable_Database_Dictionary.pdf

16. U.S. Census Bureau. American Community Survey 2012 1yr Estimates Summary File: Technical Documentation. https://www.socialexplorer.com/data/ACS2012/documentation/e348c232-c9be-4f82-9568-9c343a734e52. Accessed December 16, 2015.