# Hospital choice in a government funded health insurance scheme: Evidence from Andhra Pradesh

Tripathi, Shruti

National Institute of Public Finance and Policy

31 May 2018

# Hospital choice in a government funded health insurance scheme: Evidence from Andhra Pradesh

May 31, 2018

**Abstract**

This study examines the factors that influence patient's choice of a hospital when health-care is financed by government funded health insurance scheme. The model is estimated using a multinomial logit applied to about 0.3 million cases of inpatient treatment from one of the state health insurance scheme in India in 2015. This is the first attempt to identify and quantify the impact of individual and hospital specific factors on patient choice for tertiary care under an insurance scheme in India. The results show that in absence of price constraint patients prefer to choose providers believed to be of higher quality in our case private and big public hospitals, bypassing the smaller public hospitals.

JEL Classification: I130, I180
Key words: health insurance, patient's choice, public and private, health care financing, government policy

# 1 Introduction

The private sector over the time has become an important actor in the systems which are characterised by public (government) and private - provision of health services. In developing countries, the private sector provides a substantial proportion of health-care to the population. For example, in South Asia, about three-quarters of the children from the poorest income quintile with acute respiratory conditions seeking health-care go to a private provider (Mundial (2004)), and about 45% of sick from the poorest income quintile across 26 African countries go to a formal or informal private provider rather than a public provider for health-care (Marek *et al.* (2005)). This high utilisation of private health care in the absence of any form of health protection scheme has lead to high out of pocket expenditure by the households seeking health care. As a result, a significant proportion of households face a burden of expenditure that is catastrophic[1] for household welfare, and can lead to households falling further into poverty (Xu *et al.* (2003)).

Health systems across countries have developed specifically to allow people to use the health services they might need while protecting them against the adverse financial consequences of paying for care. A tax-based health financing mechanism, as in UK, Canada and Australia or a broad based social health insurance programs as in Germany, France, Mexico, etc. is being prescribed as a key instrument of health financing strategy for many developing countries like India (Gottret and Schieber (2006)). It has also been argued that given the failure or limitations of the public sector and high demand of private sector in developing economies, contracting of private sector may be an appropriate response to scale up the service delivery when thinking about universal health care (Paul *et al.* (2016)).

However, inclusion and high demand of private sector in the absence of efficient public sector will imply an increasing fiscal expenditure on health in medium to long run. Given the limits to government funding, this phenomenon will put a question mark on such health system's sustainability in the future. Therefore, understanding the patients' choice of health-care services and their preferences is increasingly becoming an important public policy strategy aiming to improve efficiency of the public expenditure in

[1] Catastrophic health expenditure, occurs when a household's out-of-pocket (OOP) payments are so high relative to its available resources that the household foregoes the consumption of other necessary goods and services. Impoverishment, occurs when OOP payments push households below or further below the poverty line, a threshold under which even the most basic standard of living is not ensured.

many countries across the globe (Department (2010), Lyles *et al.* (2002)).

India, in recent years has seen adoption of numerous state funded health insurance schemes with an aim to protect poor families from catastrophic inpatient health expenses. (Patnaik, Shah (2017), Ravi and Bergkvist (2013)). The main features of the schemes are: exclusive focus on tertiary care [2] hospitalization; targeted towards below poverty level population and fully funded by the state government. The study of empaneled hospitals for any of the schemes reveal that most empaneled hospitals are private facilities, reason being limited capacity in the public sector to provide tertiary care (Reddy *et al.* (2011), La Forgia and Nagpal (2012)). With the government paying for a large network of public sector health facilities and services as well as the health insurance, a dual financial burden is incurred. Therefore, recently government stressed on reinforcing trust in public health-care system by making it efficient, patient centric, affordable and effective with a comprehensive package of services that meet immediate health-care needs of poor ( National Health Policy (2017)). In order to achieve this, understanding of the factors that drive the demand for hospital care is of utmost importance. There is thus a need to study the factors that affect patient's choice for health-care in order to understand this high demand for the private care. Achieving the right mix of services from the two sectors by promoting competition driven by user's choice can substantially contribute to the qualitative development of the health systems and at the same time maintain the cost.

Evidence available from the analyses of users' revealed preference, that is, where actual patient choices have been observed, show that patients' choices depend on factors including structure (the availability of providers, the accessibility of the providers, the type and size of the providers, the availability/experience/quality of the staff, the organization of healthcare), process (availability of information, continuity of treatment, waiting time and the quality of treatment) and outcome (mortality). The importance attached to the different factors differ between patients, depending on their socio-demographic and disease characteristics. Patients generally prefer clean hospitals with high-quality services and medical qualification/expertise of providers, however, the number of beds does not influence the choice of a hospital (Gauthier and Wane (2011), Dijs-Elsinga *et al.* (2010), Roh *et al.* (2008), Schnatz *et al.* (2007)). Regarding accessibility or distance, generally, patients are averse to longer travel time and prefer a provider that is

[2] Specialized consultative care, usually on referral from primary or secondary medical care personnel, by specialists form tertiary care. Specialised Intensive Care Units, advanced diagnostic support services and specialized medical personnel are the key features of tertiary health-care.

close by (Sivey (2012),Tai *et al.* (2004), Burge *et al.* (2004)), however, being willing to travel negatively influence the importance attached to distance (Propper *et al.* (2007)). There is a positive relationship between age and the importance of distance (Haynes *et al.* (2003), Exworthy and Peckham (2006), Finlayson *et al.* (1999)).

In India, the published literature investigating the demand for health-care and patient's choice is recent and limited. One of the first studies was done by Aravindan and Kunchikannan (2000) for Kerala. The study listed proximity to private hospitals, lack of access to adequate care, drugs and doctors etc. as the major reasons for the non-utilisation of public healthcare services. Similar results were found by Dalal *et al.* (2009) in the nationwide study using NFHS data. The paper found absence of proper infrastructure, doctors, poor quality of care and long waiting time as primary reasons for non utilisation of public health-care. Data analysis from the interviews of pregnant and recent mothers in Hyderabad done by Klein (2011), show that the availability of medicines, equipment and continuity of care are the most important hospital attributes. The study also show that there is a choice towards private hospitals as they offer pre-specified packages which covers everything, whereas, patients in public hospitals often have to get medicines from outside. Most recently a study done by Patrick (2017), looks into the factors influencing preference for private and public healthcare institutions in Kerala. The main findings of the paper were as follows: (i) author did not find any significant relationship between the patient demography and preference for private health-care services and (ii) bad quality of care is the major deterrent in utilisation of public facility.

This paper examines the choice of a hospital for inpatient tertiary care among patients in the state of Andhra Pradesh. The insurance scheme of Andhra Pradesh, N.T.R. Vaidya Seva, is one of the most successful state government funded health insurance schemes in India (Reddy *et al.* (2011), La Forgia and Nagpal (2012)). We use patient level data from the NTR Vaidya Seva insurance scheme website, to model the decision to visit any hospital made by the individuals in AP when they are seriously ill and require specialised treatment. We analyse a multinomial logit model to estimate hospital demand conditional on hospital characteristics, allowing for heterogeneity across patients.

The question that this paper imposes is: What are the factors explaining the patient's choice to visit a public or private hospitals under the NTRVS scheme of Andhra Pradesh? We aim to explain the current higher usage of private sector and in turn enable governments to improve public health

services accordingly so as to bring down the fiscal costs in the future. We find that all the patients across each caste and gender choose to go to big government hospitals over private hospitals. However, when there is a choice between small government hospitals and private hospitals the odds of choosing private hospital is higher. This result implies that patients don't go to community health centers (4 doctors, 50 beds) or area hospitals (7-9 doctors, $\leq$ 100 beds), but they prefer visiting district hospitals or teaching hospitals over private hospitals. This may be because smaller public hospitals are secondary in nature with very few tertiary care facilities and the scheme only covers illness of tertiary nature. Since, large public hospitals are few in number with limited bed capacity and long waiting time, patients are forced to choose private hospitals over public.
Age seems to have very small impact, in favour of private hospitals. As the number of diseases for which hospitals are empaneled increases, the preference towards public hospital increases. Distance traveled and hospital efficiency do not have any impact on hospital demand. Our work differ from other studies in several important ways. First, our assessment is empirical in nature rather than analytical. Second, for the first time in our knowledge, hospitalisation data from any scheme is evaluated to understand patient's choice. Third, we calculate the distance between the hospital and the village of the patient to understand the impact of distance on service utilisation. Fourth, we also analyse machine learning models to understand the importance of variables in explaining the hospital demand. The rest of the article is organised as follows: section 2 summarises the scheme and section 3 explains the model of hospital choice and methodology adopted. Section 4 describes the data, section 5 discusses the results of the demand model estimations. Finally, section 6 concludes.

## 2 The N.T.R.Vaidya Seva Scheme

The NTR Vaidya Seva scheme was introduced in 2007 in Andhra Pradesh as Rajeev Arogyashree. It was renamed as N.T.R. Vaidya Seva(NTRVS) Scheme after the state's bifurcation in 2014. It targets poor households; however, due to Andhra Pradesh's (A.P.) high poverty line [3], the state government claims that in practice most of the population is covered. The scheme is operated through a trust called N.T.R Vaidya Seva Trust. The

[3] According to the state's Food and civil supplies department, BPL population consist of households with annual income not exceeding Rs.75,000 in urban and Rs.60,000 in rural areas

trust is responsible for the functioning of the scheme on both demand and supply side; from enrollment of the beneficiaries to the medical treatment and from empanelment of hospitals to their claim settlement. The scheme focuses on hospital care, and largely on tertiary hospital care. It provides coverage for the 1044 "Listed Therapies" for identified diseases in the 29 categories through a network of empanelled hospitals. The scheme empanels both public (government) and private hospitals. However, number of empaneled private hospitals are twice than public hospitals. As shown in Figure 1, in 10 districts of Andhra Pradesh private sector have significantly higher bed density than the public sector. In only 3 districts viz., Chittoor, Kadapa and Srikakulam public sector hospitals have higher density. The scheme levies no co-payments and relies entirely on general revenues (at the state level) for its finances. It provides coverage for the services to the beneficiaries up to Rs.0.25 Mn per family per annum on a floater basis[4]. As a result of the expansion of the scheme from just 20% of the population in 2008 to 85% of the population in 2014 % [5], the total expenditure on the scheme has also increased. Figure 2 shows that the claim amount[6] has doubled since 2010 from Rs.5 Bn in 2010 to little more than Rs.10 Bn in 2017.

As the scheme pays everything in the required threshold, it is purely the choice of the beneficiary that is reflected in her visit.

# 3 Model of hospital choice

We model the demand for hospitals by a multinomial logit model, following McFadden (1974). In the context of the current study , an individual can choose among three alternatives: treatment from private hospital, treatment from small public hospital and treatment from large public hospital. An individual chooses among alternatives based on the utility derived from each alternative. The utility of choice j to individual i is $U_{ij}$:

$$U_{ij} = V_{ij}(H_j, Z_j) + \epsilon_{ij} \tag{1}$$

where V(H,Z) represents utility determined by observed data.
H is a vector of hospital characteristics

[4]In a family floater insurance scheme, the benefit can be utilised by any of member for any number of time till the monetary limit is reached

[5]Source: `http://www.ntrvaidyaseva.ap.gov.in/web/guest/ntrvs`

[6]The sum paid to the empaneled hospitals after a patient, who is availing any treatment covered under the scheme, is discharged

**Figure 1** Hospital bed density defined as number of beds per 10,000 individuals. 10 districts of Andhra Pradesh have significantly higher bed density in the private sector than the public sector. In 3 districts viz. Chittoor, Kadapa and Srikakulam public sector hospitals have higher density.

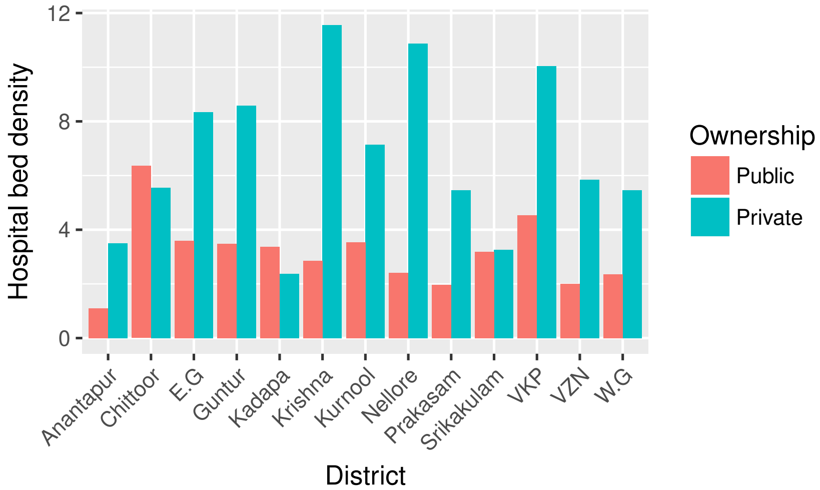
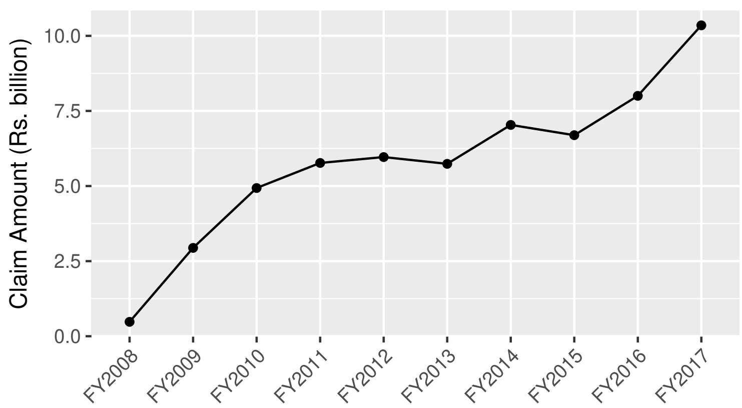



**Figure 2** Total claim amount (in Rs. Bn). Past one decade has seen a continuous increase in the claim amount. Claim amount has doubled since 2010 from 5 Bn in 2010 to little more than 10 Bn in 2017



Z is a vector of individual characteristics
$\epsilon$ is an error associated, assumed to be a random noise.
j denotes choice alternatives of hospitals depending on their ownership (0= private hospital, 1 = small public hospital, 2 = large public hospital).

Utility maximisation implies that individual i will only choose a particular alternative j if $U_{ij} > U_{ik}$, for all k not equal to j. Since $\epsilon$ is assumed to be random, $U_{ij} > U_{ik}$ is also a random occurrence. The probability of any given alternative j being chosen by an individual can be expressed as:

$$P_j = P(U_{ij} > U_{ik}) for all k \neq j \tag{2}$$

By substitution of equation 1 in 2, we get:

$$P_j = P(V_{ij} + \epsilon_{ij} > V_{ik} + \epsilon_{ik}) \forall k \neq j \tag{3}$$

Rearranging the above equation, we get:

$$P_j = P(\epsilon_{ij} - \epsilon_{ik} > V_{ik} - V_{ij}) \forall k \neq j \tag{4}$$

By using the above equation we can calculate the probability that the patient will choose alternative j. Where the $\epsilon_{ij}$ are independently and identically distributed (i.i.d.) according to the type one extreme value distribution, the probability that a patient will choose hospital j is:

$$P_j = P(V_{ij} > V_{ik}) = \frac{exp(v_{ij})}{\Sigma_{k=0}^{2} exp(v_{ik})} \forall k \neq j \tag{5}$$

We assume the observable portion of utility is a linear function of variables and coefficients $V_{ij} = x_{ij}\beta$, therefore, we operationalise equation 1 as:

$$U_{ij} = x_{ij}\beta_j + \epsilon_{ij} \tag{6}$$

Where $\beta_j$ is a vector of coefficient values indicating the effect of the vector of characteristics on an individual's utility for a hospital choice. The most widely used qualitative choice model in the literature is logistic regression. Since the patients alternative choices are more than one, we chose a multinomial logit model for the analysis. We estimate equation 6 to calculate 5 as:

$$Prob(Hosp_j|X_{ij}) = \frac{exp(x_{ij}\beta_j}{\Sigma_{k=0}^{2} exp(x_{ik}\beta_k)} \quad (7)$$

The parameters of this model can be estimated using maximum-likelihood methods. An alternative to the multinomial logit model is the nested logit model. However, since all our right hand side variables are individual characteristics, the nested logit model will essentially produce the same results as the multinomial logit model.

# 4 Data

The analysis in the paper is based on the N.T.R Vaidya Seva Scheme's patient discharge database for the year 2015. Data has information for every individual who was discharged from hospital after receiving treatment under the scheme. For each patient, information is available on patient's demography like their age, sex, caste, place of residence and details about their ailment like whether it is a medical or surgical requirement, procedure followed, time of admission, discharge and the outcome of the procedure undertaken in every case. The data also gives information about the hospital at which the patient was treated including hospital's ownership, amount of pre-authorisation[7] and final claim amount settled. We also calculate the distance between patients and the hospital they visited as the straight line distance (in kilometers) from each patient's village to the exact location of each hospital [8]

Figure 3 shows the proportion of beneficiaries availing the services from the private providers by different individual characteristics. Data shows that (1) utilisation of private health-care services increases with age; (2) female children are taken to public facilities maximum number of times; (3) general caste go to private facilities 78% of the times, followed by backward class[9]

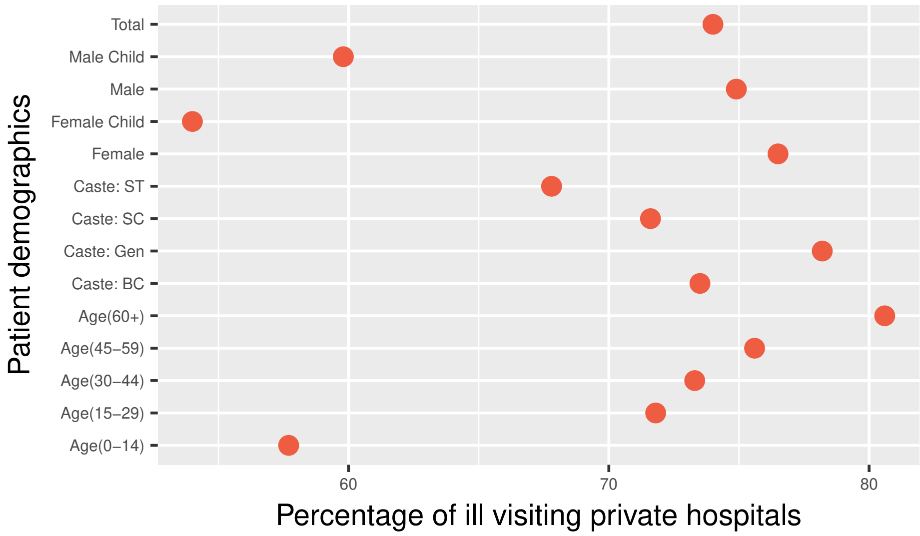
The density plot of variable distance traveled (Figure 4) shows that there is a willingness to travel among the patients. Mean distance traveled by patients

[7]or provisional authorisation, where the insurer only communicates to the hospital whether the claim is admissible or not.

[8]We were unable to find distance between patients and the hospitals in 20,581 cases. As a result our sample fell from 3,88,987 members who received treatment under the scheme to 3,62,556.

[9]In India, caste system is a form of social stratification. Highest in the caste hierarchy are general/open caste. Scheduled castes and tribes, are the lowest in the caste ranking. OBCs (Other Backward Classes) fall between the traditional upper castes and the lowest

**Figure 3** Percent of patients utilising private providers by selected characteristics. For example, 78 % of Patients falling under general category and 81 % patients over 60 years of age went to private hospital




to the empaneled hospital was at 137.2 kms and median distance traveled, 71 kms. Maximum distance traveled was quite high (850 kms), which can be explained as few patients from interior A.P traveling to a particular city where their family members may live to get a treatment.

In the year 2015, there were only 162 government hospitals empaneled under the scheme as compared to 305 private hospitals. By plotting the location of all 467 hospital on A.P. map (Figure 5), we found that most of the hospitals are located in the coastal areas, forcing patients from districts like Anantpur and Kadapa (white spaces on map) to travel. A district's utilisation of any type of health-care is also dependent on its availability. By looking at number of beds as an indicator for availability of health-care services we tried to explore the link between the the service available and its usage (Table 1). We found that in districts like Chittoor and Kadapa, where number of beds in public sector are higher than the number of beds in private sector, utilisation of services by public sector is also relatively higher. About 57% and 40% of the ill in the districts of Chittoor and Kadapa visit public health-care facilities, respectively. On the other hand, in districts like Nellore, where cumulative number of beds in public hospital are very low (714 in public and 3224 in private), utilisation of private care is as high as 90%.

The scheme covers 1044 procedures, which are pooled together in 29 cate-

**Figure 4** Kernel density plot for 'distance': The plot shows that mean distance traveled by patients in Andhra Pradesh in 2015 was about 71 kms and maximum distance traveled was 800 kms.

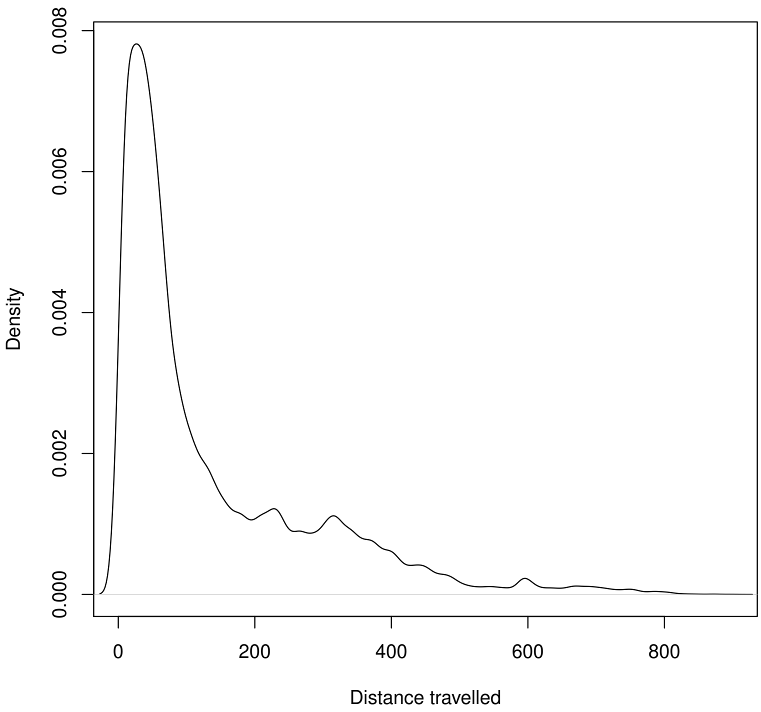



**Figure 5** Location of public and private hospitals in Andhra Pradesh

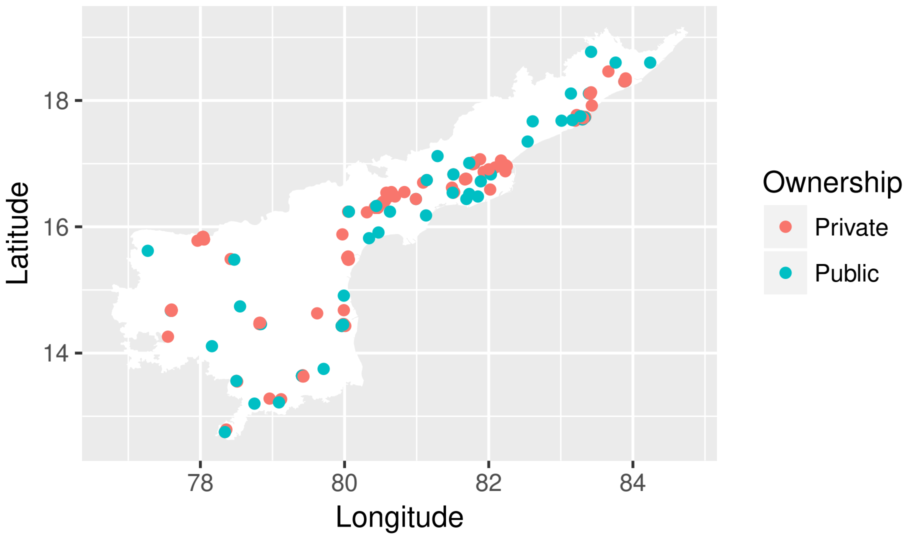


**Table 1** District wise total number of empaneled hospitals and total number of beds in the hospitals by the ownership: private and government. Table also provides district-wise information on patients utilising private providers. While in most of the district majority of the populaiton go to private hospitals, in Chittoor only 43 percent of population go to private hospitals.

| District | Count | | Tot. number of beds | | Patients in Pvt hosp. |
|---|---|---|---|---|---|
| | Pub. | Pvt. | Pub. | Pvt. | % of tot. |
| Srikakulam | 7 | 6 | 860 | 881 | 77.6 |
| Vizianagaram | 14 | 10 | 470 | 1370 | 78.6 |
| Visakhapatnam | 13 | 38 | 1947 | 4309 | 68.6 |
| East Godavari | 23 | 41 | 1895 | 4407 | 77.3 |
| West Godavari | 21 | 12 | 940 | 2182 | 83.9 |
| Krishna | 7 | 36 | 1288 | 5218 | 84.3 |
| Guntur | 13 | 38 | 1700 | 4194 | 72.6 |
| Prakasam | 9 | 20 | 670 | 1856 | 78.5 |
| Nellore | 6 | 19 | 714 | 3224 | 90 |
| Kadapa | 7 | 12 | 970 | 686 | 57.5 |
| Kurnool | 9 | 27 | 1431 | 2896 | 61.7 |
| Anantapuram | 12 | 21 | 450 | 1431 | 77.1 |
| Chittoor | 21 | 25 | 2653 | 2320 | 43.4 |
| **AP** | 162 | 305 | 15988 | 34974 | 74.2 |

gories and these categories can be further classified into surgical(invasive) and medical(non-invasive) in nature (Table 2). When we look at the utilisation of public and private hospitals with respect to surgical and medical procedures, we find that the patients went to private hospitals more for surgical procedures than medical procedure. In the case of dermatology, rheumatology and to a great extent general medicine and endocrinology related medical procedures, patients went to public hospitals.

Overall, the data set reflects that beneficiaries covered under the scheme prefer receiving services from private hospitals than public/government hospitals. The public sector's share in total beds in empaneled hospital is 32%, however it attracts only 26% of the patients. This under utilisation of public sector along with the expansion of the scheme will impose increasing fiscal burden on the government in coming years. Therefore, understanding the underlying factors that drive the choice between public and private hospitals for the patients can enable governments to improve public health services, thus bringing down the fiscal cost.

**Table 2** Medical and surgical procedures in private hospitals: Out of 13 major categories of medical procedures 8 categories shows higher utilisation of private hospitals and out of 14 categories of surgical procedures 12 categories show higher utilisation of private hospitals.

| Surgeries | 2015 | Medical | 2015 |
|---|---|---|---|
| Cardiac and Cardiothoracic Surgery | 90 | Cardiology | 87 |
| Cochlear Implant Surgery | 94 | Critical Care | 82 |
| Ent Surgery | 72 | Radiation Oncology | 81 |
| Poly Trauma | 89 | Pulmonology | 73 |
| Genito Urinary Surgeries | 96 | Gastroenterology | 62 |
| Surgical Oncology | 82 | Neurology | 53 |
| Neurosurgery | 68 | Medical Oncology | 76 |
| Ophthalmology Surgery | 95 | Nephrology | 59 |
| Orthopedic Surgery and Procedures | 93 | General Medicine | 23 |
| Pediatric Surgeries | 62 | Pediatrics | 47 |
| Plastic Surgery | 54 | Endocrinology | 28 |
| General Surgery | 53 | Dermatology | 0 |
| Surgical Gastro Enterology | 41 | Rheumatology | 0 |
| Gynaecology and Obstetrics Surgery | 40 | | |

# 5 Results

Table 3 describes the variables and their summary statistics. The results of the MNL regression estimation are presented in Table 4. In our model on the right hand side we have age, relevance, distance and efficiency as continuous variables and sex, caste and procedure are discrete (Table 3). The dependent variable is the kind of hospital that a patient visits, on the basis of ownership. We have divided it into three categories; private, small public(60 beds or less, covers CHC and area hospitals) and big public hospitals (more than 60 beds). Along with the coefficients and their significance, we also present relative risk ratio (RRR) for each provider kind, given a particular characteristic. RRR, defined as exponentiated value of the logit coefficients, allows for their easier interpretation. It can be interpreted as the relative probability of choosing a small or large public hospital relative to a private hospital conditional on patients demographics hospital characteristics.

Our main results are as follows:

1. Patient demographics: All the patients across each caste and gender choose to go to big government hospitals over private hospitals. Though, when there is a choice between small government hospitals and private hospitals, the odds of choosing private hospital is higher. We can infer that, given the access of services patients prefer big pub-

**Table 3** Description of variables

| Variable | Description | Mean(S.D.) |
|---|---|---|
| Hosp-cat (dependent variable) | 0 if private hospital, 1 if small public hospital and 2 of large public hospital | 0.51(0.86) |
| Age | Integer from 0 to 105 | 43.74(18.73) |
| Sex | Category variable: Male, Female, Male(child), Female(child) | |
| Caste | Category variable: General, Scheduled caste, Scheduled tribes, Other backward classes, Minorities | |
| Relevance | Number of specialties for which hospitals are empaneled in the scheme | 15.7(9.7) |
| Distance | Distance in km between the village of patient to the hospital visited | 134.7(146.3) |
| Efficiency | Average number of hours between pre-authorisation approval and start of surgery (hrs) | 35(44) |
| Procedure | Category variable: procedure for which patient admitted in the hospital: Surgical or Medical | |

**Table 4** Multinomial logit results

| Variable | Small public Hospitals | | Large public Hospitals | |
|---|---|---|---|---|
| | Coeff (Std. Err.) | RRR | Coeff (Std. Err.) | RRR |
| Age | -0.04*** (0.001) | 0.96 | -0.01 (0.00) | 0.99 |
| SexFemale(Child) | 0.26*** (0.00) | 1.3 | 1.14*** (0.001) | 3.13 |
| SexMale | -0.94*** (0.001) | 0.39 | 0.20*** (0.014) | 1.22 |
| SexMale(Child) | -0.30*** (0.00) | 0.74 | 1.03*** (0.001) | 2.8 |
| CasteMinorities | -0.87*** (0.00) | 0.42 | 0.20*** (0.003) | 1.22 |
| CasteOC | -0.25*** (0.001) | 0.78 | 0.05*** (0.017) | 1.05 |
| CasteOthers | -1.30*** (0.00) | 0.27 | 0.376*** (0.00) | 1.46 |
| CasteSC | 0.14*** (0.002) | 1.15 | 0.25*** (0.018) | 1.29 |
| CasteST | -0.02*** (0.00) | 0.98 | 0.47*** (0.002) | 1.6 |
| Relevance | 0.17*** (0.001) | 1.18 | 0.56*** (0.001) | 1.75 |
| Distance | 0.00 (0.00) | 1.00 | 0.00 (0.00) | 1.00 |
| Efficiency | 0.00*** (0.00) | 1.00 | 0.00 (0.00) | 1.00 |
| Procedure | 0.66 (0.001) | 1.9 | -1.37*** (0.014) | 0.3 |

RRR: Relative Risk Ratio

lic hospitals, private hospitals and small public hospitals in the given order. Age seems to have very small impact on patient's decision, in favour of private hospitals. In contrast with the literature (Sivey (2012), Beckert *et al.* (2012), Propper *et al.* (2007)), we find that distance does not affect a patient's choice. It may be due to the fact that the packages offered under the scheme for any procedure include a component for transport. Patients are paid upto Rs 500 for any cost incurred towards transportation at the time of discharge.

2. Hospital characteristics: As the number of disease for which hospitals are empaneled (relevance) increases, the preference towards public hospital increases. In contrast with other studies (Tai *et al.* (2004), Gowrisankaran and Town (2003)) efficiency - that is the time lapse between acceptance of pre-authorisation and start of the surgery does not have any impact on hospital choice. This implies that patients are ready to incur implicit cost in the form of longer waiting time. The result of hospital preference for surgical or medical procedures suggest that when given a choice between private and small public hospital, the odds are in favor of choosing private hospital for surgical procedures. There is preference for small government hospital in case medical procedures. On the other hand when the beneficiaries get to choose between big public hospitals and private hospitals then they prefer government hospital for surgery and private hospital for medical procedures.

## 5.1 Machine Learning

As a robustness check and to analyse the prediction power of logit model, we estimated the supervised machine learning models. Machine learning is the field of study that gives computers the ability to learn without being explicitly programmed. Machine learning algorithms are often categorized as supervised or unsupervised. In supervised learning, we have a data set and already know what our correct output should look like, having the idea that there is a relationship between the input and the output. Under machine learning we have esitmated two algorithms: Random Forests (bagging) and Gradient Boosting Machines(GBM) (boosting)[10] to compare and understand whether the model (estimated through logit) is a good representation of the variation in choice present in the data. Unlike logistic regression, where a

[10]are a machine learning ensemble meta-algorithm designed to improve the stability and accuracy (bagging) and reducing bias and variance(boosting) of machine learning algorithms used in statistical classification and regression

**Table 5** Machine learning results

| Variables | RF: MDG | GBM: RI |
|---|---|---|
| No of Specialities | 73403.12 | 88.39 |
| Efficiency | 29181.93 | 11.55 |
| Procedure | 1905.24 | 0.049 |
| Age | 1269.26 | 0.006 |
| Distance | 929.26 | 0 |
| Sex | 898.58 | 0 |
| Caste | 590.92 | 0 |

statistical model that was likely to have generated the data is specified by the researcher prior to estimation, machine learning technique doesn't impose any such requirement at pre-estimation stage. Random forests (Breiman (2001)) are an ensemble method used for classification and regression. The methodology includes construction of decision trees created by using bootstrap samples of the training data and random feature selection in tree induction. Prediction is made by aggregating (majority vote or averaging) the predictions of the ensemble created by using test data. When using gradient boosting technique (Friedman, 2001), choice of algorithm to create classification or solve regression problem is with the researcher. It can be decision tree, neural network, support vector machine or any other such algorithm. It builds the model in a stage-wise fashion. The subsequent samples depend on weights given to records in the previous sample which did not predict correctly. The final prediction is also not a simple average of all the predictions made, but a weighted average. Following common practice, we divide data into training and testing data in a 70:30 ratio. We train a GBM and random forest of 500 decision trees, each of them on $\sqrt{p}$ predictors, where p is the total number of predictors.
These techniques make no assumptions about the distribution of the data and are less prone to over-fitting of the data than a traditional regression techniques. Therefore, resulting in an improvement over traditional paramteric fitting of model.

Machine learning techniques do not provide results like standard regression and therefore, they lack interpretability. But what they lack in interpretation, they more than make up for in prediction power. Also they can produce a list of predictor variables that they believe to be important in predicting the outcome. They are used to rank the importance of independent variables. Random forests uses mean decrease in Gini to give out the list. The mean decrease in Gini coefficient is a measure of how each variable contributes to the homogeneity of the nodes in the resulting random forest. Each time a particular variable is used to split a node, the Gini coefficient for the

**Table 6** Performance matrix-comparing three models

| | Private | Small Pub | Large Pub |
|---|---|---|---|
| Recall | | | |
| Random Forest | 1 | 0.75 | 0.99 |
| GBM | 0.85 | 1 | 0.99 |
| Logit | 0.94 | - | 0.84 |
| Precision | | | |
| Random Forest | 0.99 | 1 | 0.99 |
| GBM | 0.94 | - | 0.97 |
| Logit | 0.93 | - | 0.82 |

child nodes are calculated and compared to that of the original node. The changes in Gini are summed for each variable and normalized at the end of the calculation. Variables that result in nodes with higher purity have a higher decrease in Gini coefficient. GBM estimate the relative influence of predictor variables. The measures are based on the number of times a variable is selected for splitting, weighted by the squared improvement to the model as a result of each split, and averaged over all trees. Table 5 presents each variable in the data and their importance as calculated by the two methods. They are ordered top-to -bottom as most to least important. Results show that hospital related variables like number of specialties for which hospitals are empanelled, efficiency and whether going for surgical or medical procedure are the factors on which choice of hospitals depend.

## 5.2 Performance matrix

In this section, we summarise the performance of the three models. Performance metrics to evaluate discrete classification problems (Powers, 2011) are accuracy, precision, recall and F-score. To estimate the accuracy of a test, we calculate the proportion of true positive and true negative in all evaluated cases. Precision and recall measure the number of true positives relative to false positives (type I error) and false negatives (type II error), respectively. Maximum accuracy was achieved at 99% for Random Forest, followed by 95% for GBM and 90% for logit analysis. This shows that over all fit of the model estimated from logit regression is good and comparable with machine learning models. When comparing class wise prediction across models (Table 6), we see that logit predictions are comparable with Random Forests and GBM predictions. Therefore, MNL not only utilise all the important variables, as suggested by machine learning exercise, in the modeling but also has good accuracy in predicting hospital demand.

# 6 Conclusion

Our results are consistent with the patterns of utilisation by individuals covered under insurance and facing zero monetary price of medical care. In the absence of price constraint, the patients prefer to choose providers believed to be of higher quality in our case private and big public hospitals, bypassing the smaller public hospitals. Inclusion of hospital specific characteristics in hospital demand shows that patient are aware about the quality and availability of services provided by patients in their choice set. The fact that efficiency does not have a significant impact on patients' decision implies that due to low supply of high quality services, with all explicit cost covered, patients are willing to bear the implicit cost of higher waiting time.

Thus, improving quality of existing public health-care services would reduce the fiscal burden in the short run, while expanding their coverage would make this scheme fiscally sustainable in the longer run.

# References


Aravindan K, Kunchikannan T (2000). "Changes in Health status of Kerala 1987–1997." *Center For Development Studies. Thiruvananthapuram.*

Beckert W, Christensen M, Collyer K (2012). "Choice of NHS-funded Hospital Services in England." *The Economic Journal*, **122**(560), 400–417.

Breiman L (2001). "Random forests." *Machine learning*, **45**(1), 5–32.

Burge P, Devlin N, Appleby J, Rohr C, Grant J (2004). "Do patients always prefer quicker treatment?" *Applied Health Economics and Health Policy*, **3**(4), 183–194.

Dalal K, Dawad S, *et al.* (2009). "Non-utilization of public healthcare facilities: examining the reasons through a national study of women in India." *Rural Remote Health*, **9**(3), 1178.

Department OE (2010). "Health care systems: Getting more value for money." *Policy Notes No. 2.*

Dijs-Elsinga J, Otten W, Versluijs MM, Smeets HJ, Kievit J, Vree R, van der Made WJ, Marang-van de Mheen PJ (2010). "Choosing a hospital for surgery: the importance of information on quality of care." *Medical Decision Making*, **30**(5), 544–555.

Exworthy M, Peckham S (2006). "Access, choice and travel: implications for health policy." *Social Policy & Administration*, **40**(3), 267–287.

Finlayson SR, Birkmeyer JD, Tosteson AN, Nease Jr RF (1999). "Patient preferences for location of care: implications for regionalization." *Medical care*, **37**(2), 204–209.

Friedman JH (2001). "Greedy function approximation: a gradient boosting machine." *Annals of statistics*, pp. 1189–1232.

Gauthier B, Wane W (2011). "Bypassing health providers: the quest for better price and quality of health care in Chad." *Social Science & Medicine*, **73**(4), 540–549.

Gottret PE, Schieber G (2006). *Health financing revisited: a practitioner's guide.* World Bank Publications.

Gowrisankaran G, Town RJ (2003). "Competition, payers, and hospital quality." *Health services research*, **38**(6p1), 1403–1422.

Haynes R, Lovett A, Sünnenberg G (2003). "Potential accessibility, travel time, and consumer choice: geographical variations in general medical practice registrations in Eastern England." *Environment and Planning A*, **35**(10), 1733–1750.

Klein T (2011). “Why Do India’s Urban Poor Choose to Go Private? Hospital Choice Experiments in Slums of Hyderabad.” *Journal [serial on the Internet].*

La Forgia G, Nagpal S (2012). *Government-sponsored health insurance in India: Are you covered?* World Bank Publications.

Lyles A, Weiner JP, Shore AD, Christianson J, Solberg LI, Drury P (2002). “Cost and quality trends in direct contracting arrangements.” *Health Affairs*, **21**(1), 89–102.

Marek T, O’Farrell C, Yamamoto C, Zable I (2005). “Trends and opportunities in public-private partnerships to improve health service delivery in Africa.” *Africa Region Human Development Working Paper Series. Washington (District of Columbia): World Bank.*

Mundial B (2004). “World development report 2004: making services work for poor people.” *Banque mondiale, Washington, DC.*

Patrick M (2017). “Public and Private Healthcare Institutions: Preference and Expenditure Pattern.”

Paul E, Bodson O, Ridde V, Fecher F (2016). “Universal health coverage in Low and Middle income countries: Economic Analyses.” *Reflets et perspectives de la vie economique*, **55**(1).

Powers DM (2011). “Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation.”

Propper C, Damiani M, Leckie G, Dixon J (2007). “Impact of patients’ socioeconomic status on the distance travelled for hospital admission in the English National Health Service.” *Journal of Health Services Research & Policy*, **12**(3), 153–159.

Ravi S, Bergkvist S (2013). “Are Publicly Financed Health Insurance Schemes Working in India?”

Reddy KS, Selvaraj S, Rao KD, Chokshi M, Kumar P, Arora V, Bhokare S, Ganguly I (2011). “A critical assessment of the existing health insurance models in India.” *Public Health Foundation of India*, pp. 1–115.

Roh CY, Lee KH, Fottler MD (2008). “Determinants of hospital choice of rural hospital patients: the impact of networks, service scopes, and market competition.” *Journal of Medical Systems*, **32**(4), 343–353.

Schnatz PF, Murphy JL, OâĂŹSullivan DM, Sorosky JI (2007). “Patient choice: comparing criteria for selecting an obstetrician-gynecologist based on image, gender, and professional attributes.” *American Journal of Obstetrics and Gynecology*, **197**(5), 548–e1.

Sivey P (2012). "The effect of waiting time and distance on hospital choice for English cataract patients." *Health Economics*, **21**(4), 444–456.

Tai WTC, Porell FW, Adams EK (2004). "Hospital choice of rural medicare beneficiaries: patient, hospital attributes, and the patient–physician relationship." *Health services research*, **39**(6p1), 1903–1922.

Xu K, Evans DB, Kawabata K, Zeramdini R, Klavus J, Murray CJ (2003). "Household catastrophic health expenditure: a multicountry analysis." *The lancet*, **362**(9378), 111–117.