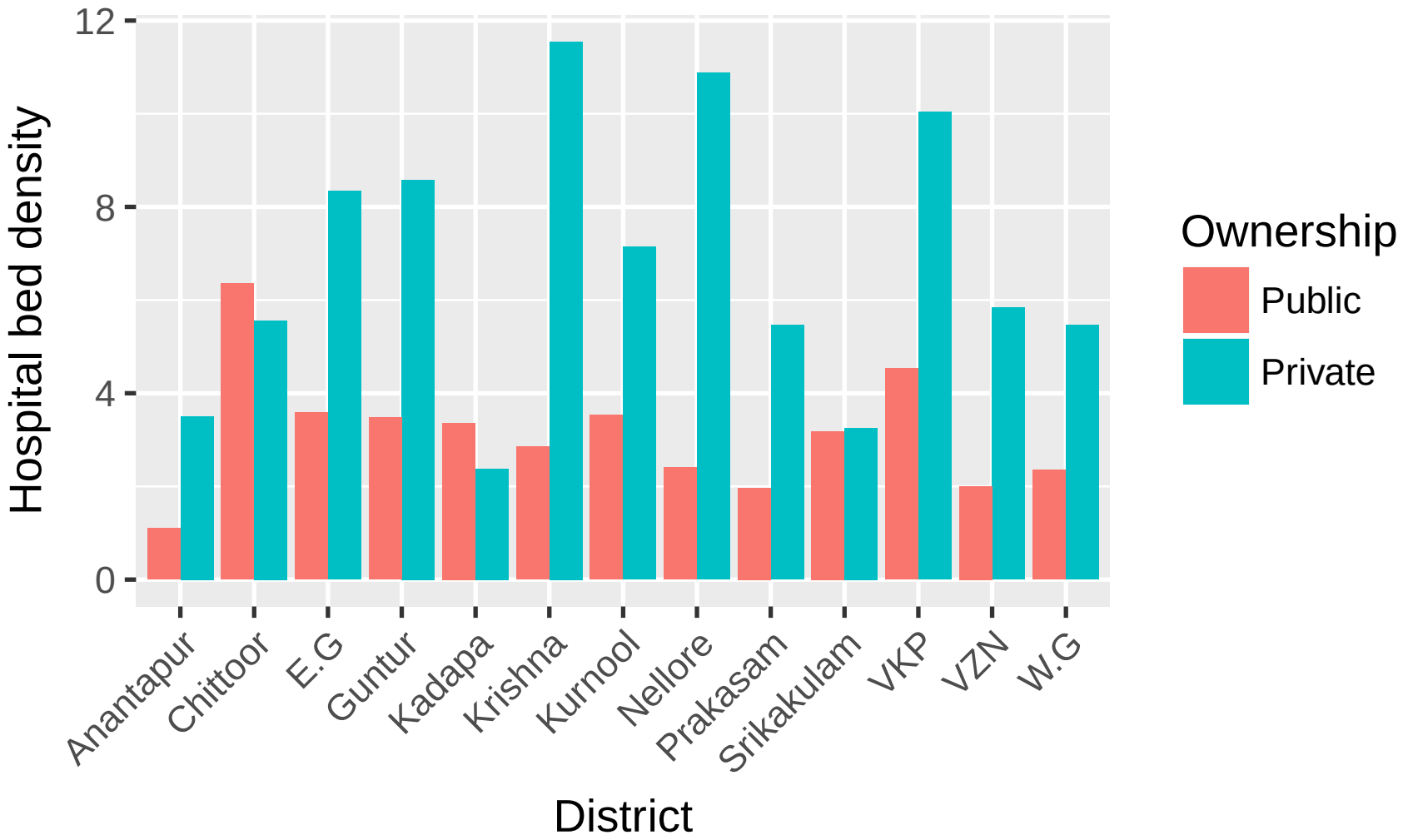
Hospital choice in a government funded health insurance scheme: Evidence from Andhra Pradesh

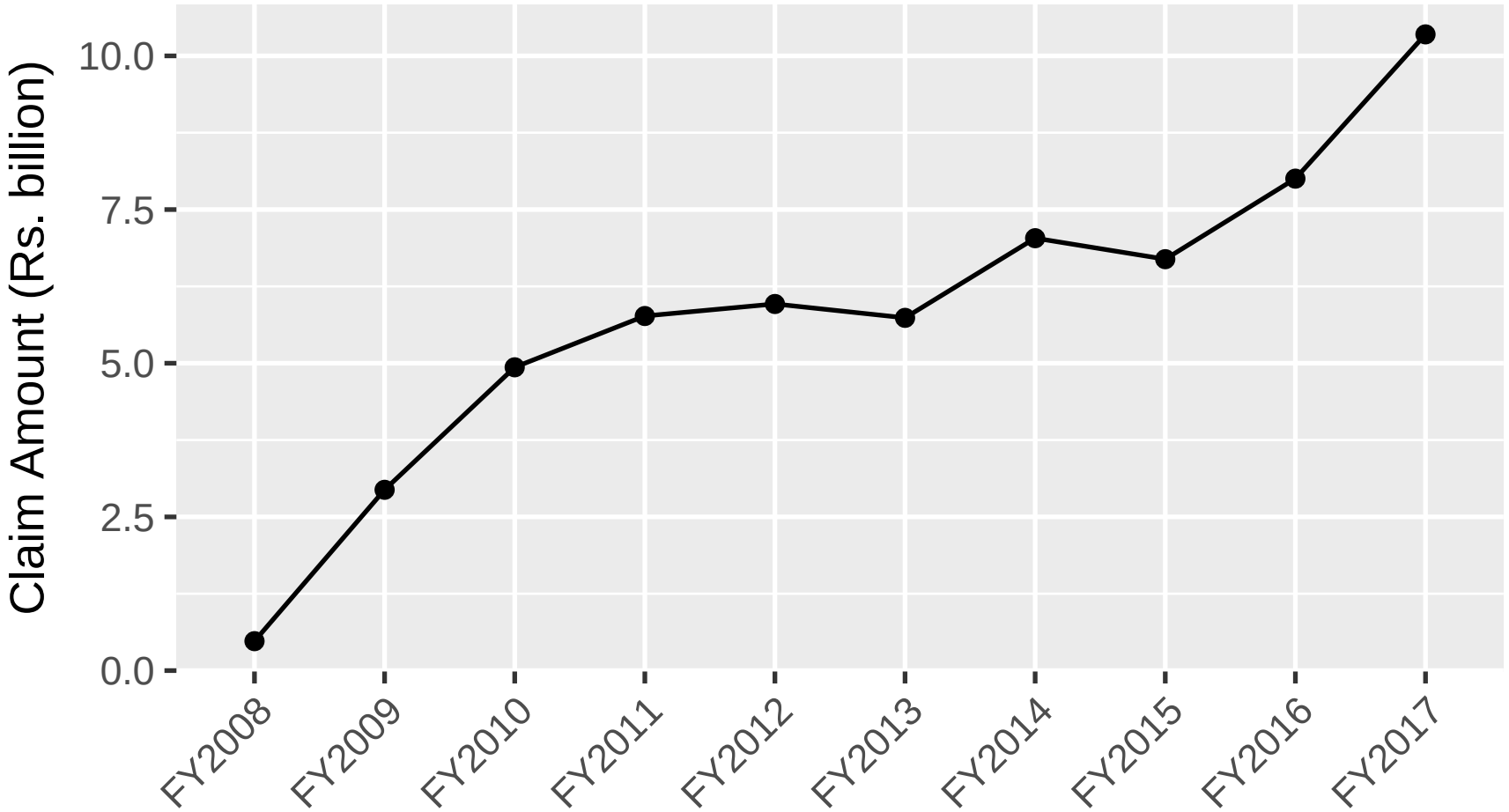
Motivation

- GFHIS were introduced with the objective to save poor families from catastrophic health expenditure
- Concentrate on Tertiary care
- Existing number of hospital beds density in public hospitals per 10,000 of the population is low.
- For instance in Andhra Pradesh: density is 3.13 beds in public hospitals as compared to 6.77 beds in private hospitals.

Density of beds in public hospitals(per 10,000 individuals)



Total claim amount paid under the scheme



Motivation

- Understand the factors that drive the choice between public and private hospitals.
- Enable governments to improve public health services, bring down the fiscal costs.

Research Questions

- What are the factors explaining the choice to visit a public or private hospitals under the NTRVS scheme of Andhra Pradesh?
- Conditional on distance, what is the preference between public and private hospitals for patients?

Scheme: Introduction

- NTR Vaidya Seva (Rajeev Arogyashree) was introduced in 2007
- Aim: to cover illnesses falling under the scope of 'Tertiary Care' which causes high OOP health expenditure by BPL families.
- Population coverage: BPL population - due to AP's high poverty line most of the population is covered.
- Financial coverage: Up to Rs. 0.25 Mn per family; per annum on floater basis. Levies no co-payments, funded through general revenues of the government.
- Procedure coverage: Provides end to end cash-less services for identified diseases (1044 listed therapies in 29 categories) under secondary and tertiary care through empaneled hospitals.
- As the scheme pays everything in the required threshold, it is purely the choice of the member that is reflected in her visit.

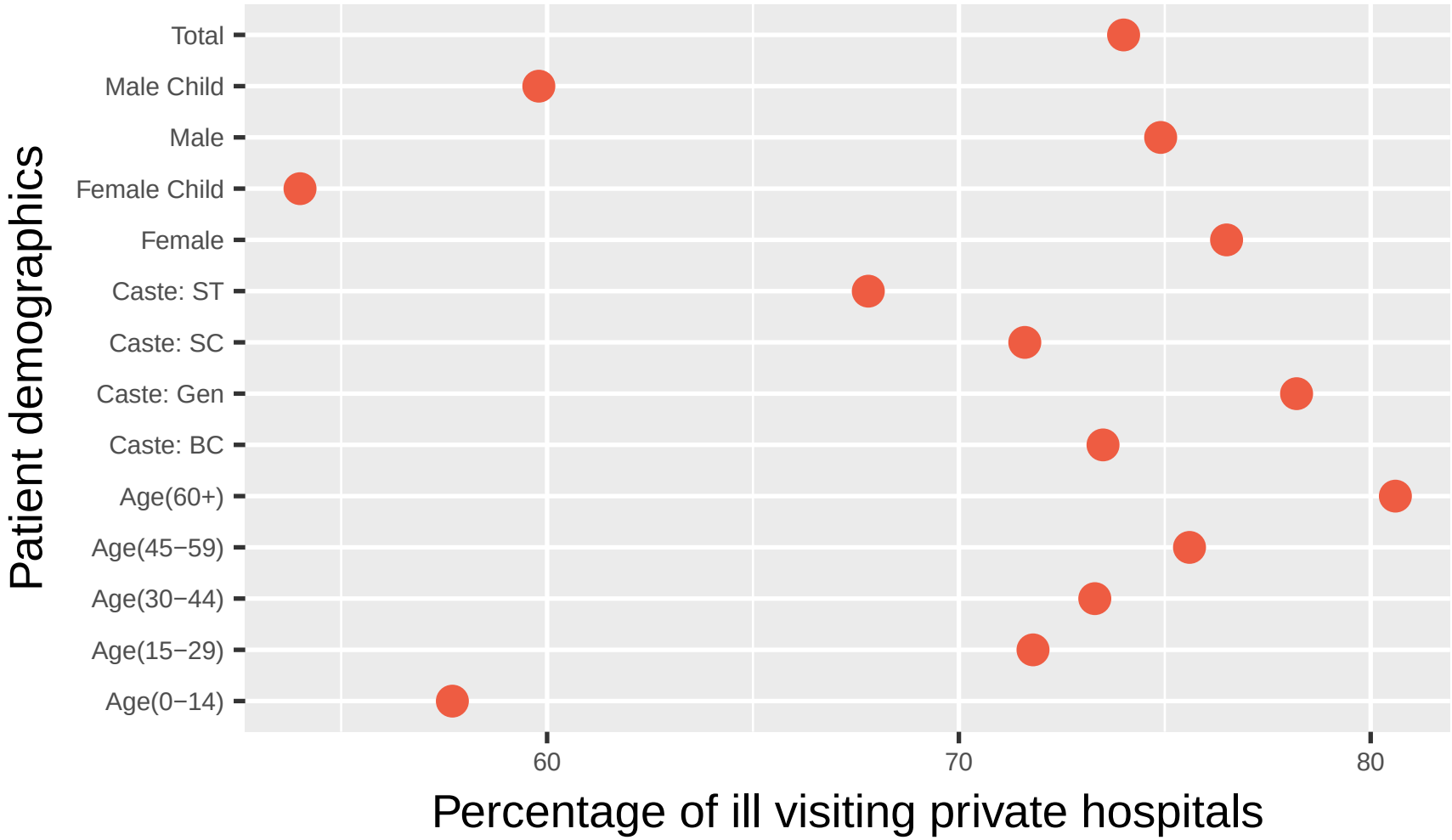
Data: Source and characteristics

- Data is extracted from NTRVS Trust Website: `http://www.ntrvaidyaseva.ap.gov.in`
- Patient-level claims data for 2015
- Variables: patient demography, hospital information, nature of illness.
- Variables: Distance from village to hospital(s) - calculated through finding the Euclidean distance between geographical coordinates of villages and empaneled hospitals.
- Total observations: 4,80,940
- Percent of ill visiting - (i) Private Hospitals - 73.66% (ii) Public Hospitals - 26.34%

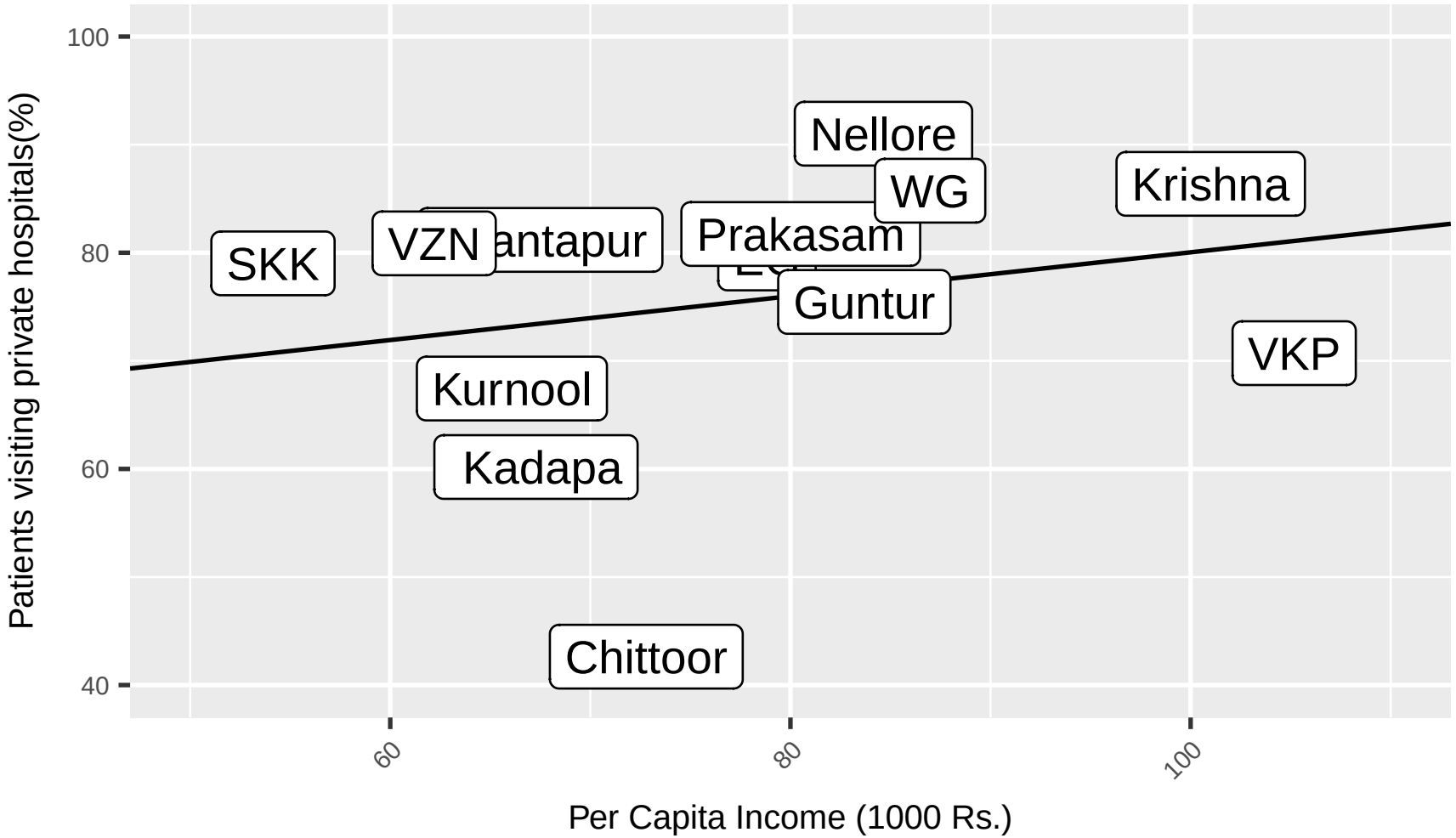
Summary Statistics
Continuous variables

Variable	Min.	Max.	Mean	Median	S.D.
Age	0	105	43.74	46	18.71
Distance(in km)	0.122	901.79	137.23	70.68	149.97
Claim Amount	500	250000	25141	20000	24962

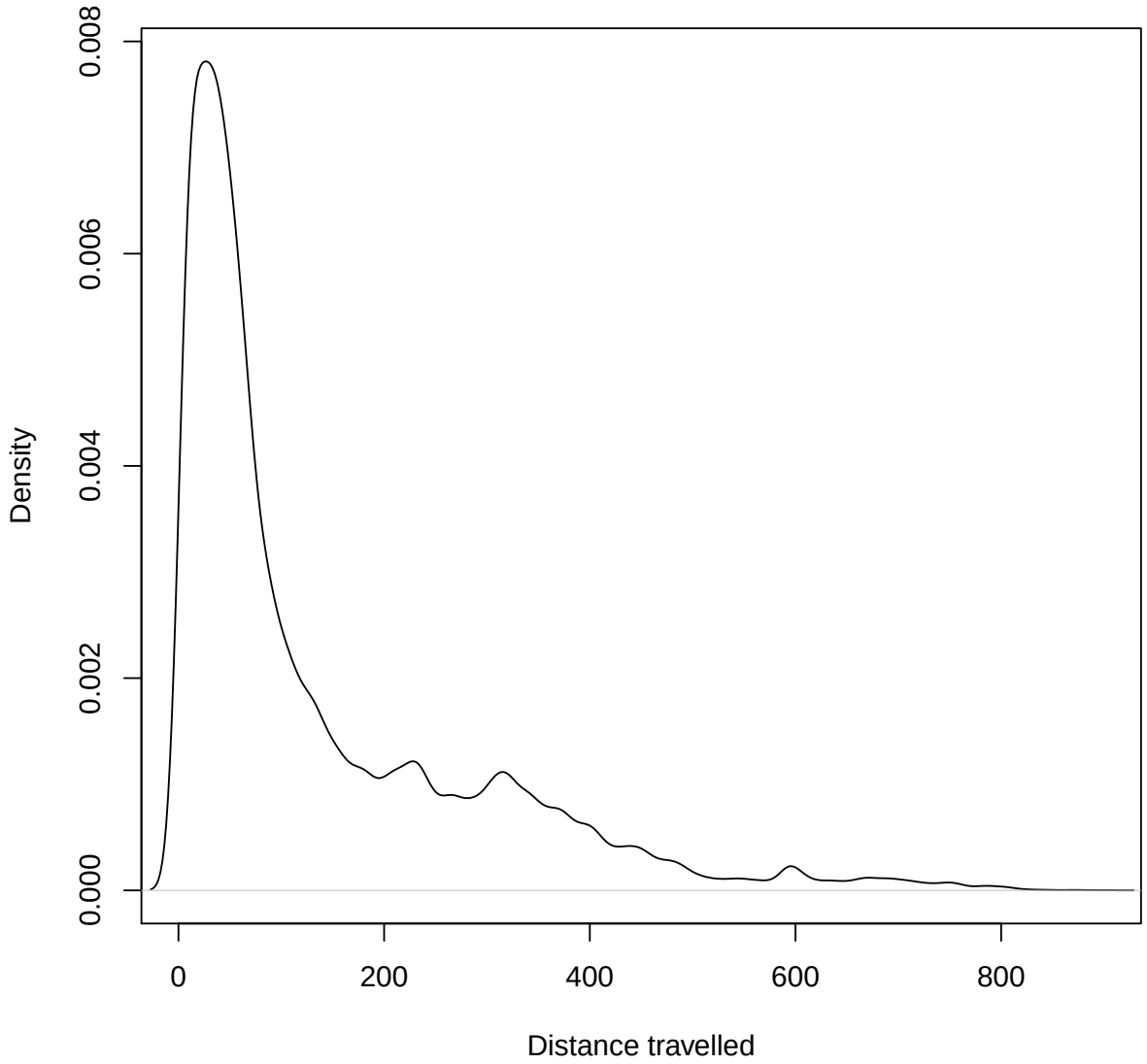
Percent of ill visiting private providers by selected characteristics: Social characteristics



Percent of ill visiting private providers by selected characteristics: Economic status



Kernel density plot for 'distance'

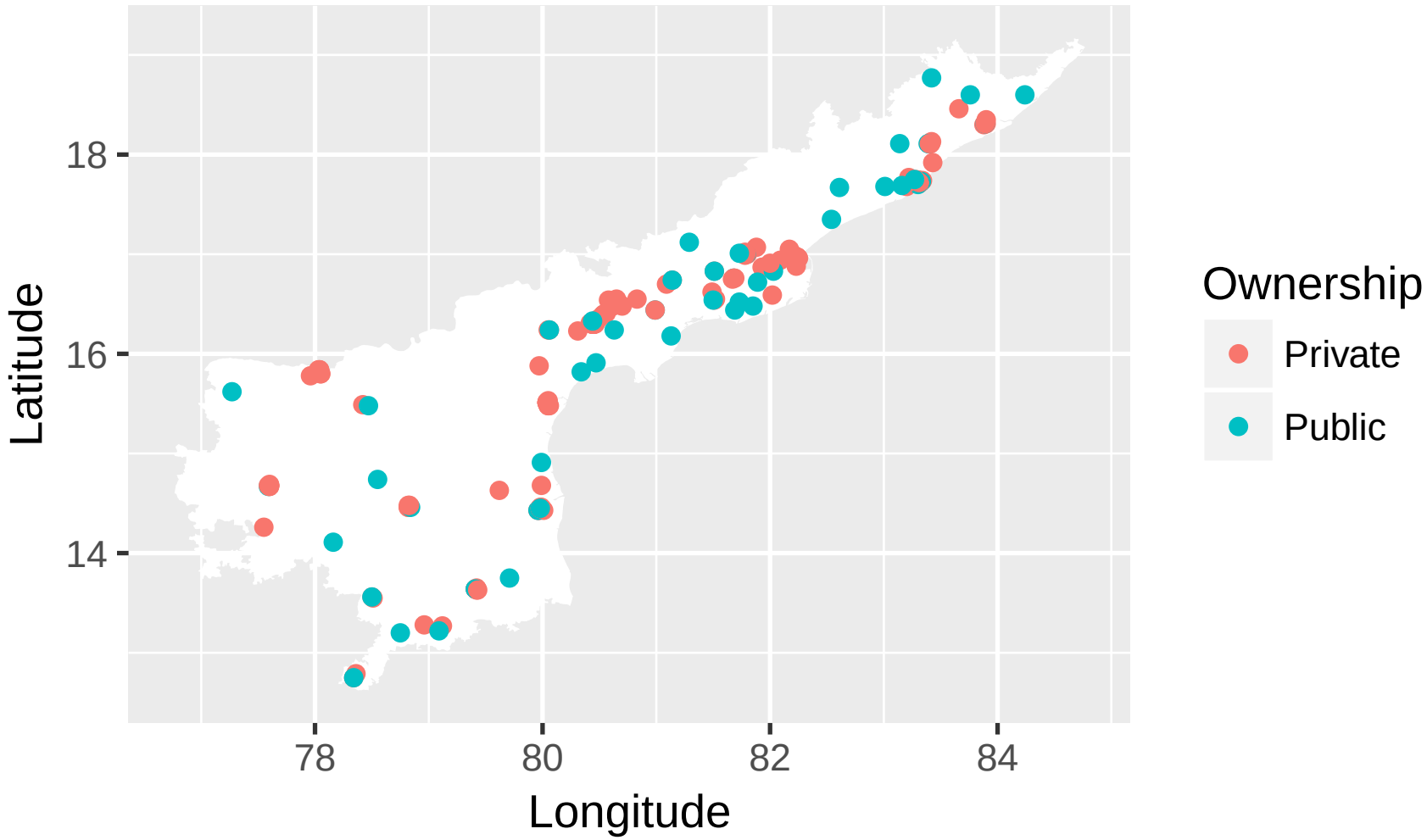


Summary Statistics: Hospital-level data

Public-Private split: Medical procedures in private hospitals

Surgeries	2015	Therapies	2015
Cardiac and Cardiothoracic Surgery	90	Cardiology	87
Cochlear Implant Surgery	94	Critical Care	82
Ent Surgery	72	Radiation Oncology	81
General Surgery	53	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
Poly Trauma	89	Dermatology	0
Surgical Gastro Enterology	41	Rheumatology	0
Gynaecology and Obstetrics Surgery	40		

Distribution of empaneled public and private hospitals in Andhra Pradesh



Possible factors driving an individual/household's healthcare service utilisation choice

- An individual's use of health services is considered to be a function of three categories of characteristics[1]:
  - Predisposing factors: social structure (education, ethnicity,culture), health beliefs (attitudes, values), demographic (age, gender)
  - Enabling factors: personal (income, health insurance, travel), community (available facilities, waiting time)
  - Need factors: perceived (individual's view of their own health), evaluated (professional judgement)

[1] **AN1995**

Question I

- Question: Factors determining health service utilisation - is distance/ease of accessibility relevant?
- Dependent variable/label: dummy variable 0(visit to private hospital), 1(visit to public hospital)

Logistic Regression Result - AP

Dependent variable: Private hospital = 0, Public hospital = 1

	OR	2.5%	97.5%	p-value
No.of.Specialties	0.98	0.98	0.98	0.00
dist	1.00	1.00	1.00	0.04
SorTtherapy	1.37	1.36	1.39	0.00
CasteOC	0.90	0.89	0.91	0.00
CasteMinorities	1.05	1.02	1.07	0.00
CasteOthers	1.16	0.95	1.41	0.13
CasteSC	1.11	1.09	1.12	0.00
CasteST	1.24	1.20	1.27	0.00
Age	0.99	0.99	0.99	0.00
SexFemale(Child)	1.37	1.32	1.42	0.00
SexMale	1.04	1.03	1.05	0.00
SexMale(Child)	1.14	1.11	1.18	0.00
DistChittoor	2.71	2.62	2.79	0.00
DistEast Godavari	1.07	1.04	1.10	0.00
DistGuntur	1.17	1.14	1.21	0.00
DistKrishna	0.71	0.69	0.73	0.00
DistKurnool	1.53	1.48	1.58	0.00
DistNellore	0.58	0.56	0.60	0.00
DistPrakasam	0.88	0.86	0.91	0.00
DistSrikakulam	0.94	0.91	0.97	0.00
DistVishakhapatnam	1.23	1.19	1.27	0.00
DistVizianagaram	0.80	0.77	0.83	0.00
DistWest Godavari	0.65	0.63	0.67	0.00
DistYSR Kadapa	2.15	2.07	2.22	0.00

Predictive analysis: Choice between public vs. private hospitals

- Along with Logit, Random Forest and Gradient-boosting models are trained on a subset of the data and then tested on the hold-out sample (70-30 split).
- The prediction metrics are far superior for Random Forest and GBM than Logit.
- Better performance can be attributed to non-linearities and multiple variable interactions that tree-based and boosting models account for.
- Performance metrics:

Variable	GBM	RF	Logit
Accuracy	91.86%	95.81%	76%
Precision	78.86%	86.96%	52%
F-score	86.02	92.2	62

- Thus, the variable importance metrics derived from the ML models need to be carefully looked at.

Machine Learning Model Results

- GBM: No. of trees = 5000, Variable interactions tried = 3
- Random Forest: No. of trees = 500, Variables tried at each split = 2
- Importance Matrix:

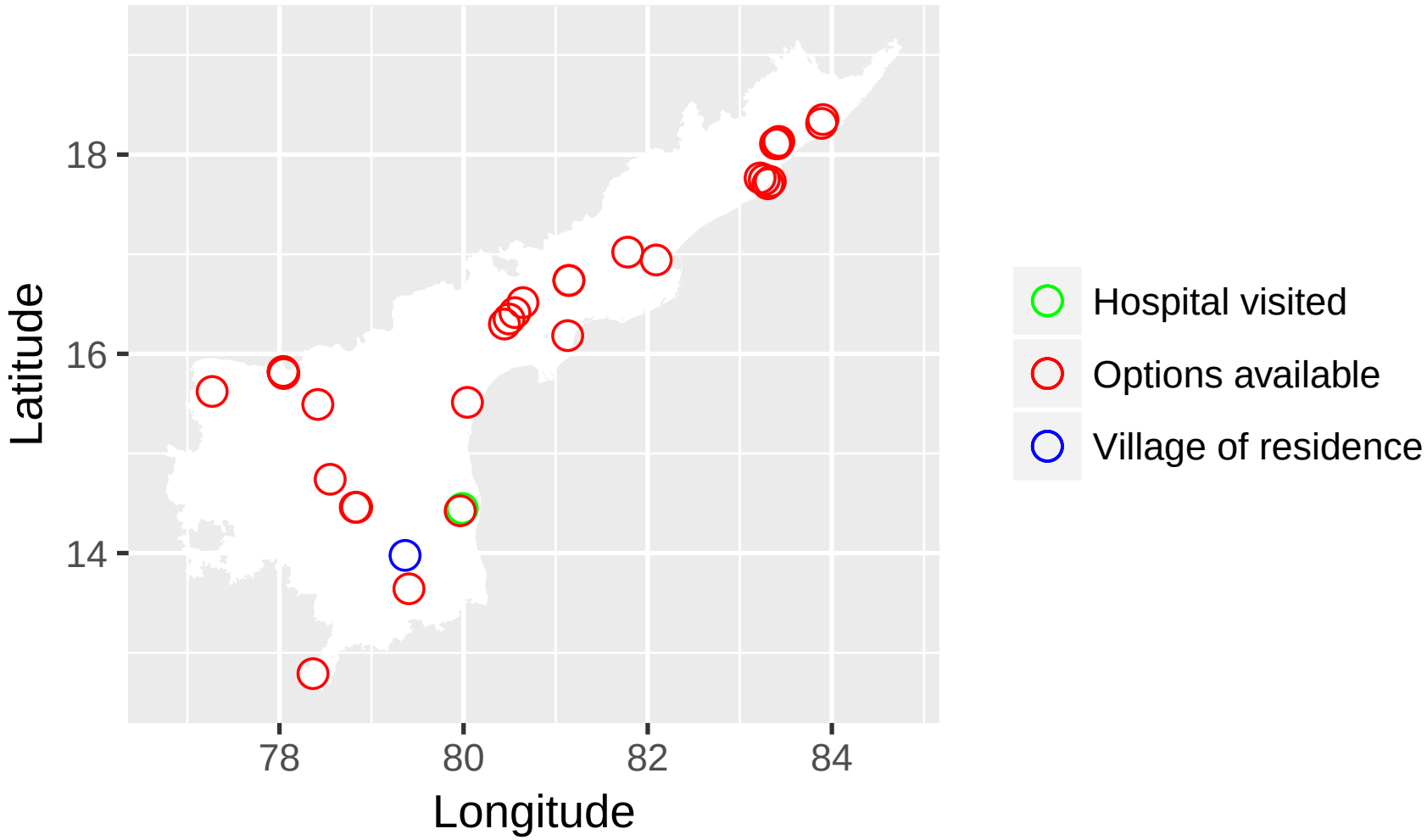
Variable	GBM: RI	RF:MDG
No. of specialties	85.72	50719.01
SorT(Therapy vs. Surgery)	9.57	13421.57
District	4.58	10685.52
Claim Amount	0.14	3287.39
Caste	0.0001	636.61
Age	0.00	1387.29
Gender	0.00	393.27
Distance(in km)	0.00	1594.41

- The models show that the size of hospital, nature of illness(Therapy or Surgery), and community-level factors(proxied by District) are the most important factors while deciding between private and public hospitals.

Question II

- Distance travelled by the patient encompasses both implicit and explicit costs.
- Question: Given the a set of alternatives in the neighborhood based on distance, is there a preference for private hospitals?

Example: Patient residence, choice set of hospitals and hospital visited



Distance-based categorical variable

Data and characteristics

- Nearest 25 hospitals selected for all villages based on Euclidean distances calculated.
- If the member chooses to go to the nearest hospital, category $= 1$
- If anywhere between nearest and 5th nearest, value $= 5$,...,
- Value $= 25$ - if the member chooses to visit any hospital in the neighbourhood between 20th and 25th nearest hospital.
- In total: 7 ordered categories

Ordered Logistic Regression result

Dependent variable: nearest = 1, farthest = 26

	OR	2.5%	97.5%	p-value
SorTtherapy	0.917	0.908	0.926	0.00
No of Specialties	1.005	1.0048	1.006	0.00
ownership	0.962	0.952	0.972	0.00
Age	1.000	1.000	1.000	0.00
SexFemale(Child)	0.871	0.925	0.887	0.00
SexMale	0.942	0.933	0.951	0.00
SexMale(Child)	0.893	0.871	0.916	0.00
CasteMinorities	1.017	0.996	1.038	0.00
CasteOC	1.036	1.024	1.048	0.00
CasteOthers	1.163	0.983	1.376	0.00
CasteSC	0.975	0.963	0.988	0.00
CasteST	1.079	1.05	1.108	0.00

Machine Learning Model Results

- GBM: No. of trees = 1000, No. of variable interactions tried = 3
- Random Forest: No. of trees = 3000, Variables tried at each split = 2
- Importance Matrix:

Variable	GBM: RI	RF:MDG
Age	4.83	1502.26
No. of Specialties	61.52	2243.56
Ownership	32.96	394.32
SorT	0.46	236.92
Caste	0.22	624.88
Gender	0.00	322.24

- The models show that the size of the hospital, age, caste of the patient, and the ownership of the hospital hold relevance for willingness to travel for the treatment.

Conclusions

- There is awareness about the scheme amongst the targeted population.
- Distance does not affect the odds of choosing private or public hospitals.
- Patients prefer to visit public hospitals for therapies relative to surgeries.
- Size of hospital, nature of illness, and place of residence have strongest influence on choice between public and private hospitals.
- Patients prefer to travel for private hospitals.
- Patients' willingness to travel for surgical procedures is greater relative to therapies.
- Age, nature of illness, hospital size and ownership have the highest impact on willingness to travel.

Thank you.