Internet use and health care utilization

Agne Suziedelyte

School of Economics
The University of New South Wales

Presentation for 7th Summer Workshop in Health Economics
January 28, 2010

Some numbers

Internet users		
	U.S.	Australia
2000	46%	47%
2008	74%	74%

e-health seekers, U.S.	
2000	25%
2008	61%

Around 60 % of e-health seekers say that their recent search had an impact on how they take care of their own or somebody else health.

Does Internet use for health affect health care utilization?

- 22% of e-health seekers say that their recent search has affected decision to see a doctor.
- The question is a topic of interest in medical and communication literature.
- No research has been done in economics literature.

Why is this research important?

- It answers the questions about the direction of relationship.
- It can assist health care policy planning.
- It may change physicians'attitude towards online health information.

Two ways to treat online health information

1. As an input into health production function:
   - online information may substitute for information from a doctor;
   - it may also increase demand for information or services from a doctor.
2. As an exogenous factor, which affects marginal product of formal medical care.

Econometric model

- Health care utilization is measured by doctor visits.
- As it is non-negative discrete variable, which takes few values, Poisson count data model is used.
- Conditional mean function is assumed to be exponential:

$$E(HC|I, \mathbf{x}) = \exp(\delta I + \mathbf{x}\beta), \ \textit{where} \tag{1}$$

HC is health care,
I is an indicator that an individual looks for health information online, and
x are controls.

Health Information National Trends Survey

- The survey is managed by the US Department of Health and Human Services.
- Its main purpose is to determine how Americans find and use health information.
- The data set contains information on information seeking behavior, health knowledge, risk perceptions and behaviors, health and personal characteristics.
- The data has been collected from January 2008 through May 2008.
- The full data set contains 7,674 observations and 473 variables.
- 5,529 of observations could be used for the analysis.

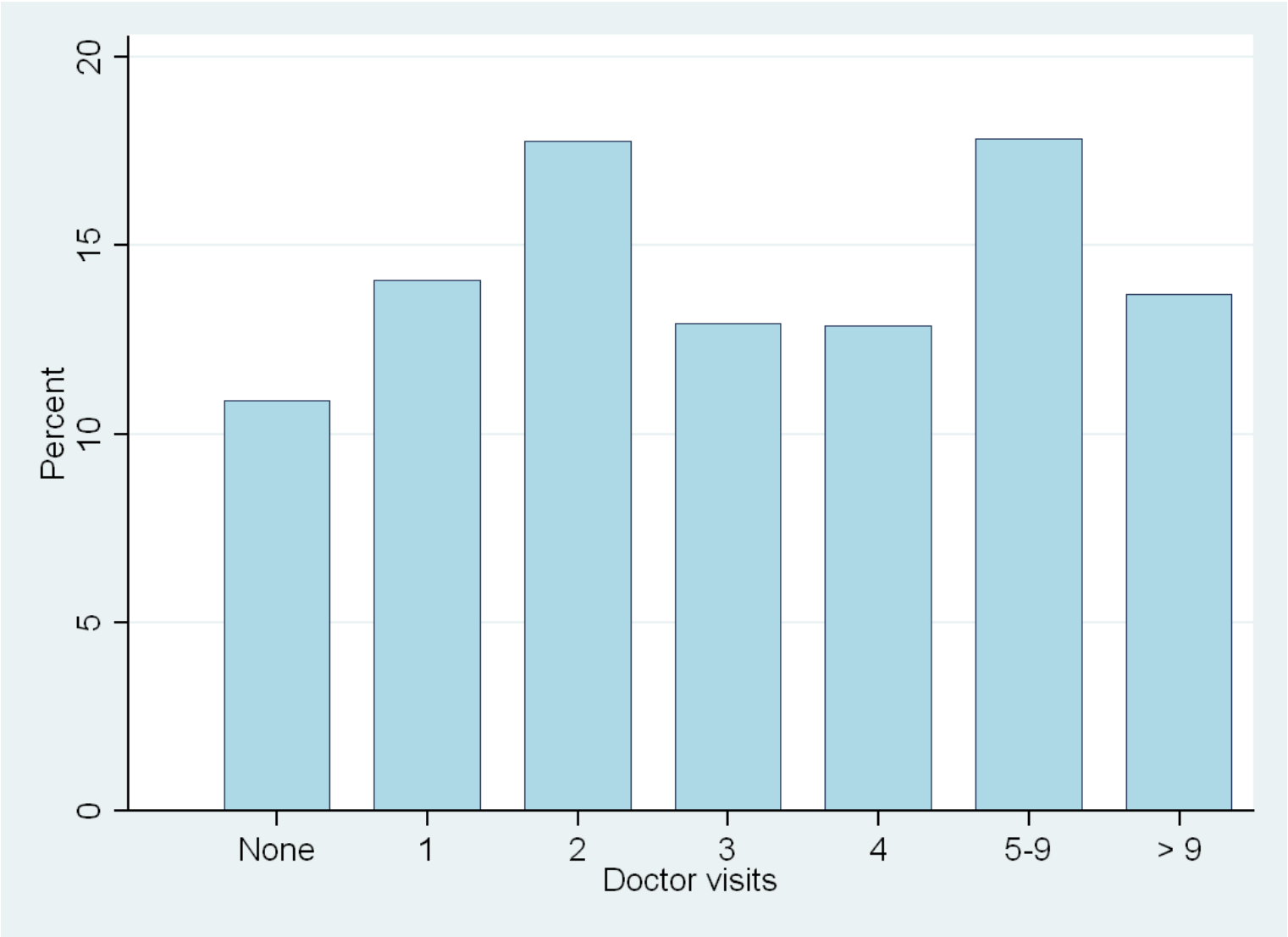
Internet use for health purposes variable

- *eHealthSeeker* is an indicator variable taking a value of 1 if an individual has used Internet in the most recent search for health information for him/herself, and 0 otherwise.
- 40 % of the sample had used Internet in their recent health information search.

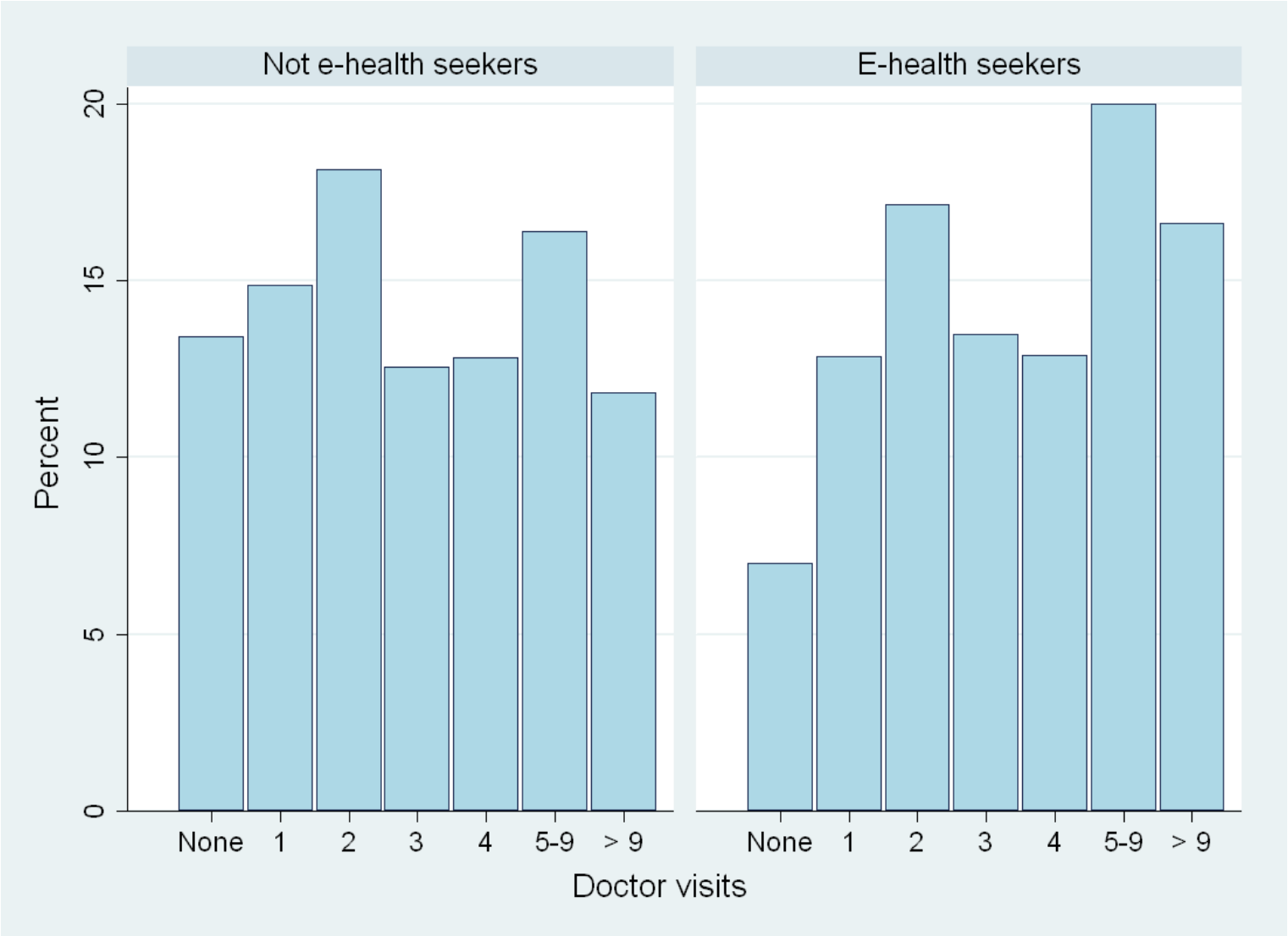
Health care utilization variable

- Doctor visits variable is an answer to the question *"In the past 12 months, not counting times you went to an emergency room, how many times did you go to a doctor, nurse, or other health professional to get care for yourself?"*
- Up to 4 visits, the actual number of visits is provided.
- If there were 5 or more visits, information whether it was '5-9' or '10 or more times'.
- The mid-point of the first interval was taken.
- A value of 12 was chosen as a mid-point of the second interval.
- Mean number of doctor visits is 4.29 (st. dev. 3.76)

Distribution of doctor visits



Distribution of doctor visits for e-health seekers and non-seekers



Controls

- The controls include measures of need (health status), ability to pay (income, insurance coverage), relationship with a doctor, and socio-demographic factors (age, sex, education, income, race, marital status, and employment status).
- Most of the people have a regular doctor, health cover, and quite a lot of trust in their physician.
- Most of the sample assess their health as good or very good.
- Around 14% have had cancer.
- The average age is 53 and 40 % of the sample are males.
- Most of the respondents are married, white, have college education, are employed, and live in a household with USD75,000 income or more in smaller metro areas.

Possible sources of endogeneity

Why may Internet use for health purposes be endogenous?

- Unobserved heterogeneity,
- Reverse causality.

Instruments

- The literature does not suggest any instruments for Internet use.
- Requirements for an instrument:
  - relevance,
  - exogeneity.
- High-speed internet (broadband or wireless) availability is an important factor of e-health seeking behavior.
- Literature on suggests that some state-level Internet policies affect high-speed Internet deployment.
- 3 types of policies regulating:
  - internet service provision by municipalities;
  - rights-of-way access by telecommunication companies;
  - financial support to broadband providers.

Information collection

- Most of the information was not readily available and needed to be collected from the primary sources.
- As a result, the information collection process was very time consuming.
- The collected information was used to construct 19 state-level dummies describing 3 policies.
- 2 were selected as the instruments:
  1. a dummy for states, which facilitate rights-of-way access and
  2. a dummy for states, which have procedures for solving disputes related to rights-of-way.

Validity of instruments - relevance

eHealthSeeker	Reduced form regression
rowfacilit	0.063***
	(0.017)
rowdispreg	-0.049**
	(0.021)
N	5547
R^2	0.14

*** significant at 1% level, ** significant at 5% level, * significant at 10% level.

- The policies are individually and jointly significant (F statistic = 9.43).
- States with policies facilitating rights-of-way access have higher fraction of e-health seekers.
- Having dispute resolution procedures in place, has a negative effect.

Validity of instruments - exogeneity

- To address possible concerns about exogeneity of the instruments, the states with and without policies were compared according to:
  - the fraction of urban population,
  - medical expenditures for physician services (by state of provider and residence),
  - total medical expenditures (by state of provider and residence).
- The mean comparison tests did not reveal any significant differences between the two state groups.

Endogeneity test results

- The results of endogeneity test show no evidence against exogeneity of Internet use for health purposes.

Doctor Visits	
$\hat{v}*$	0.067
s.e.	0.336

* $\hat{v}$ is a residual from the reduced from regression

Poisson results

Doctor Visits	Coefficient	s.e.	Robust s.e.
eHealthSeeker	0.202	0.014***	0.023***
HasRegDoctor	0.577	0.022***	0.042***
TrustInDoctor	0.059	0.012***	0.020***
VgoodHealth	0.168	0.025***	0.041***
GoodHealth	0.328	0.025***	0.041***
FairHealth	0.513	0.029***	0.047***
PoorHealth	0.673	0.038***	0.060***
HadCancer	0.265	0.017***	0.027***
Male	−0.149	0.015***	0.024***
Black	0.087	0.023***	0.036**
AmerIndian	0.050	0.055	0.086
OtherRace	−0.134	0.031***	0.053**
LessThanHighS	0.040	0.028	0.048
HighSchool	−0.058	0.019***	0.031*
SomeCollege	0.008	0.016	0.027
Unemployed	0.233	0.035***	0.064***
Homemaker	0.197	0.025***	0.043***
Student	0.132	0.046***	0.078
Retired	0.211	0.020***	0.032***
Disabled	0.491	0.028***	0.044***
HasHCoverage	0.373	0.028***	0.051***
Rural-urban code	−0.015	0.004***	0.006**
Constant	−0.071	0.067	0.109
N	5529		
Pseudo R^2	0.12		
Log likelihood	-14908.62		

*** significant at 1% level, ** significant at 5% level, * significant at 10% level.
Income, age, and marital status were also controlled for.

Interpretation of the results

- People who use Internet as health information source have around 22% more expected doctor visits than non users.
- At the mean (4.29), the effect is almost 1 visit.
- Average marginal effect = 0.88 visit.

Interpretation of the results, cont'd

	An 'average' male	A male without health cover	A male with health problems
Predicted visits	3.857	1.15	9.42
Marginal effect	.70***	.21***	1.72***
s.e.	(.09)	(.02)	(.15)

*** significant at 1% level

What do the results mean?

- The results indicate that using Internet as health information source has a positive effect on doctor visits.
- The results are consistent with the hypothesis that online information is a complement to doctor information and/ or services.
- They also support the hypothesis that doctor visits for more informed individuals are more productive.

Implications for the health care policy

- Assuming that extra visits induced by online health information are productive, it would be recommended to encourage Internet use for health and increase health information content on the Internet.
- If the case is opposite, then it is recommended to allocate resources to improving information evaluation skills and control of quality of the online health information.
- To determine the nature of the additional visits, the research needs to be extended.

Remaining tasks

- To conduct additional sensitivity analyzes (coding of doctor visits, coding of eHealthSeeker, and missing data) and confirm that the survey is nationally representative.

Endogeneity test procedure

- Let c be unobserved characteristics of an individual correlated to Internet use for health purposes:

$$E(HC|I, \mathbf{x}, c) = \exp(\delta I + \mathbf{x}\beta + c) \tag{2}$$

- Let $\mathbf{z}$ be a vector of instrumental variables, not correlated with c
- The steps of endogeneity test are:
  1. regress I on $\mathbf{z}$ and $\mathbf{x}$ and save residuals $\hat{v}$
  2. estimate the equation (2) including $\hat{v}$ as a regressor.
  3. test of null hypothesis that coefficient on $\hat{v}$ is not statistically different from zero.

IV estimation results

	Poisson	IV Poisson
eHealthSeeker	0.202***	0.214
	(0.023)	(0.830)
N	5529	5529

*** significant at 1% level, ** significant at 5% level, * significant at 10% level.

OLS and 2SLS results

	OLS	2SLS
eHealthSeeker	0.856***	0.823
	(0.099)	(1.665)
N	5529	5529

*** significant at 1% level, ** significant at 5% level, * significant at 10% level.

Other count models' results

	Poisson	NegBin I	NegBin II	ZIP
DoctorVisits1				
eHealthSeeker	0.202***	0.198***	0.219***	0.190***
	(0.023)	(0.022)	(0.024)	(0.015)
lndelta				
_cons		0.501***		
		(0.034)		
lnalpha				
_cons			-0.950***	
			(0.033)	
N	5529	5529	5529	5529
Log likelihood	-14908.62	-13137.00	-13275.36	-14764.06

*** significant at 1% level, ** significant at 5% level, * significant at 10% level.

Does the effect differ across population groups?

	pois	pois_regd	pois_noregd
eHealthSeeker	0.202***	0.191***	0.351***
	(0.023)	(0.024)	(0.083)
N	5529	4380	1149
Pseudo R^2	0.12	0.07	0.13
Log likelihood	-14908.62	-12116.39	-2659.39

	pois	pois_fph	pois_notfph
eHealthSeeker	0.202***	0.099*	0.229***
	(0.023)	(0.048)	(0.026)
N	5529	840	4689
Pseudo R^2	0.12	0.13	0.10
Log likelihood	-14908.62	-2401.24	-12447.67

*** significant at 1% level, ** significant at 5% level, * significant at 10% level.

Does the effect differ across population groups?, cont'd

	pois	pois_older	pois_younger
eHealthSeeker	0.202***	0.145***	0.279***
	(0.023)	(0.029)	(0.039)
N	5529	3155	2374
Pseudo R^2	0.12	0.11	0.13
Log likelihood	-14908.62	-8610.30	-6194.77

	pois	pois_male	pois_fem
eHealthSeeker	0.202***	0.230***	0.185***
	(0.023)	(0.039)	(0.028)
N	5529	2198	3331
Pseudo R^2	0.12	0.15	0.10
Log likelihood	-14908.62	-5682.22	-9150.18

*** significant at 1% level, ** significant at 5% level, * significant at 10% level.

Does the effect differ across population groups?, cont'd

	pois	pois_hinc	pois_linc
eHealthSeeker	0.202***	0.215***	0.199***
	(0.023)	(0.038)	(0.029)
N	5529	1875	3654
Pseudo R^2	0.12	0.09	0.14
Log likelihood	-14908.62	-5028.44	-9828.88

*** significant at 1% level, ** significant at 5% level, * significant at 10% level.