Lecture 17:
Policy Lab
Impacts of health insurance expansion I

PPHA 34600

Prof. Fiona Burlig

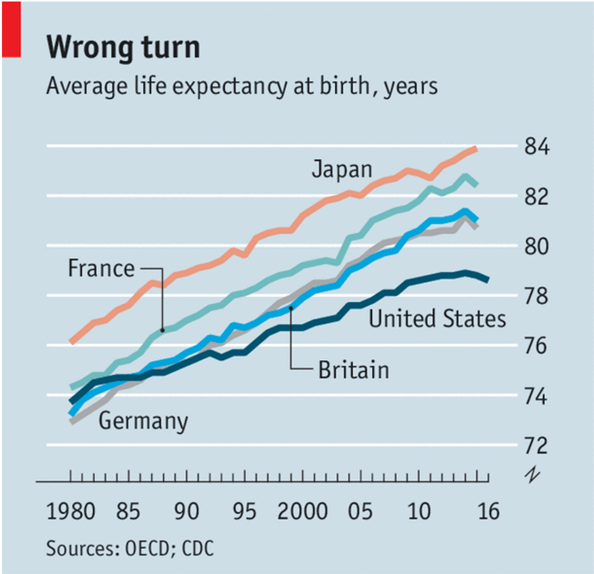
Harris School of Public Policy
University of Chicago

From last time: Machine learning for causal inference

We looked at several ways to incorporate ML into CI

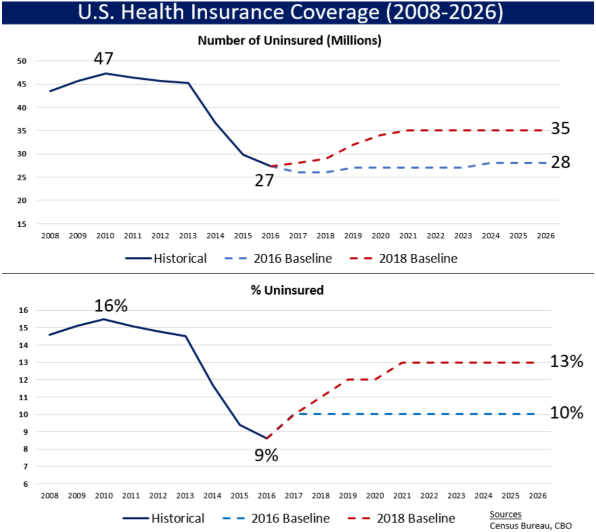
1. Generating (big) data
2. Exploring heterogeneity
3. Improving research designs
   - ML works with SOO to handle functional form
   - And with SOU to aid in generating counterfactuals

US life expectancy is low



Economist.com

As is health insurance access



Health insurance expansion is a central policy topic

OBAMACARE

Research question

What is the causal effect of health insurance on health?

Research question

What is the causal effect of health insurance on health?

This is not totally straightforward to answer:

- Naive estimator: compare insured to uninsured people

Research question

What is the causal effect of health insurance on health?

This is not totally straightforward to answer:

- Naive estimator: compare insured to uninsured people
- **Why is this problematic?**

Research question

What is the causal effect of health insurance on health?

This is not totally straightforward to answer:

- Naive estimator: compare insured to uninsured people
- **Why is this problematic?**
- Insured people might be...:
  - Wealthier (poorer) than uninsured people
  - More (less) healthy
  - More (less) employed
  - Etc

Research question

What is the causal effect of health insurance on health?

This is not totally straightforward to answer:

- Naive estimator: compare insured to uninsured people
- **Why is this problematic?**
- Insured people might be...:
  - Wealthier (poorer) than uninsured people
  - More (less) healthy
  - More (less) employed
  - Etc

→ There are many forms of selection bias!

The ideal experiment

What would we do if we could do anything?

- Some kind of random assignment to health insurance
- … but even this is not straightforward

What do we mean by "health insurance"?

What do we mean by "health insurance"?

The ideal experiment

What would we do if we could do anything?

- Some kind of random assignment to health insurance
- ... but even this is not straightforward

Important considerations:

- Do we want to randomly assign free medical care?
  - Or should actual insurance reflect existing conditions?

The ideal experiment

What would we do if we could do anything?

- Some kind of random assignment to health insurance
- ... but even this is not straightforward

Important considerations:

- Do we want to randomly assign free medical care?
  - Or should actual insurance reflect existing conditions?
- Do we care about effects on people? Hospitals? Cities? Counties?

The ideal experiment

What would we do if we could do anything?

- Some kind of random assignment to health insurance
- ... but even this is not straightforward

Important considerations:

- Do we want to randomly assign free medical care?
  - Or should actual insurance reflect existing conditions?
- Do we care about effects on people? Hospitals? Cities? Counties?
- Do we think about healthcare in one location at a time?
  - Or do we consider general equilibrium?

The ideal experiment

What would we do if we could do anything?

- Some kind of random assignment to health insurance
- ... but even this is not straightforward

Important considerations:

- Do we want to randomly assign free medical care?
  - Or should actual insurance reflect existing conditions?
- Do we care about effects on people? Hospitals? Cities? Counties?
- Do we think about healthcare in one location at a time?
  - Or do we consider general equilibrium?

→ Actually ideal experiment probably requires multiple countries

  → Or at least a really large sample!

If you thought the *ideal* experiment was tricky...

An additional practical wrinkle:

- Randomizing insurance access is impractical!
  - Very hard to randomize a government program
  - Simple randomization is therefore not really going to work
  - Not to mention that it's going to be extremely hard to randomize at any meaningful scale

If you thought the *ideal* experiment was tricky...

An additional practical wrinkle:

- Randomizing insurance access is impractical!
  - Very hard to randomize a government program
  - Simple randomization is therefore not really going to work
  - Not to mention that it's going to be extremely hard to randomize at any meaningful scale

→ A quasi-experimental approach may be useful here

First paper: Currie and Gruber (1996)

An early prominent econ paper estimating effects of healthcare expansion:

- **Research question:** What is the effect of Medicaid expansion on insurance, care utilization, and health

→ AKA, what is the effect of insurance on health?

Currie and Gruber (1996): Context

C&G study the US in the 1980s and 1990s:

- Over 30% of poor children lack any health insurance

Currie and Gruber (1996): Context

C&G study the US in the 1980s and 1990s:

- Over 30% of poor children lack any health insurance
- Doctors often don't treat publicly insured patients

Currie and Gruber (1996): Context

C&G study the US in the 1980s and 1990s:

- Over 30% of poor children lack any health insurance
- Doctors often don't treat publicly insured patients
- Medicaid limited to low-income women
  - Single-parent children eligible too

Currie and Gruber (1996): Context

C&G study the US in the 1980s and 1990s:

- Over 30% of poor children lack any health insurance
- Doctors often don't treat publicly insured patients
- Medicaid limited to low-income women
  - Single-parent children eligible too

→ Big push to expand Medicaid to other kids

→ Opportunity for a natural experiment!

Currie and Gruber (1996): Context

What did the Medicaid expansion look like?

- Medicaid was linked to Aid to Families with Dependent Children
- Eligibility cutoffs low $\rightarrow$ stigma prevents people from applying?
- Some states had other low-income programs to qualify for Medicaid
- Deficit Reduction Act of 1984: family structure requirements relaxed
- Income cutoffs raised; states have discretion

Currie and Gruber (1996): Eligibility changes

TABLE I
MEDICAID ELIGIBILITY AND COVERAGE

Year	% of children eligible	% of children eligible—fixed population	% of children covered
1984	16.1	16.1	13.2
1985	18.2	18.4	13.5
1986	19.0	18.9	13.8
1987	19.3	19.7	13.5
1988	18.8	20.3	12.8
1989	20.4	21.6	13.9
1990	25.7	26.2	16.5
1991	28.7	28.1	19.3
1992	31.2	30.3	20.6

Based on data from March 1985–March 1993 CPS. Column 1 shows the percent of children eligible for Medicaid in each year. Column 2 shows the percentage of the 1984 sample that would have been eligible for Medicaid in each subsequent year (holding their characteristics constant and inflating income appropriately). Column 3 gives the percentage of children actually covered in each year. Figures are from the authors' calculations as described in the text and in Appendix 1.

Currie and Gruber (1996): Data

TABLE III
NHIS SAMPLE MEANS BY MEDICAID ELIGIBILITY

	All	Medicaid eligible
% Eligible	0.219	1.000
	(0.001)	
# Observations	227,169	49,991
Utilization of care		
No doctor's visits last 12 months	0.194	0.197
	(0.001)	(0.002)
Doctor's visit last 2 weeks	.115	0.118
	(0.001)	(0.001)
Any hospitalization last 12 months	0.036	0.049
	(0.0004)	(0.001)
Visit to doctor's office last 2 weeks	0.087	0.071
	(0.0006)	(0.001)
Visit to ER or hospital clinic last 2 weeks	0.017	0.027
	(0.0003)	(0.001)
Visit to other site of care last 2 weeks	.015	0.024
	(0.0003)	(0.001)
Family & child characteristics		
Male	0.513	0.509
	(0.001)	(0.002)
Black	0.180	0.355
	(0.001)	(0.002)
Hispanic	0.120	0.219
	(0.001)	(0.002)
Age	6.873	5.401
	(0.001)	(0.019)
Female head/spouse is HS dropout	0.240	0.506
	(0.001)	(0.002)

Currie and Gruber (1996): Estimation approach

Medicaid expansions were not randomly assigned:

- We need a research design to estimate the causal effect of interest

Ideally, we'd estimate:

$$Y_i = \alpha + \tau D_i + \varepsilon_i$$

where:

Y_i is the outcome (utilization/health)
D_i is a Medicaid eligibility indicator
ε_{id} is an error term

$\rightarrow$ Without random assignment, we will get bias (why?)

Currie and Gruber (1996): Estimation approach

Without random assignment, we could leverage time:

$$Y_{it} = \tau D_{it} + \alpha_s + \delta_t + \varepsilon_{idt}$$

Currie and Gruber (1996): Estimation approach

Without random assignment, we could leverage time:

$$Y_{it} = \tau D_{it} + \alpha_s + \delta_t + \varepsilon_{idt}$$

→ Even with time, we still have identification concerns (why?)

Currie and Gruber (1996): Estimation approach

Without random assignment, we could leverage time:

$$Y_{it} = \tau D_{it} + \alpha_s + \delta_t + \varepsilon_{idt}$$

$\rightarrow$ Even with time, we still have identification concerns (why?)

- Omitted variables
- Endogeneity

Currie and Gruber (1996): Estimation approach

C&G uses an IV approach to overcome these problems:

- We want to isolate the effect of Medicaid from everything else

For the instrument to be valid, we need:

1. **First stage:** Our IV needs to be correlated with Medicaid eligibility
2. **Exclusion restriction:** Our IV needs to only move Y through eligibility

Currie and Gruber (1996): Estimation approach

C&G uses an IV approach to overcome these problems:

- We want to isolate the effect of Medicaid from everything else

For the instrument to be valid, we need:

1. **First stage:** Our IV needs to be correlated with Medicaid eligibility
2. **Exclusion restriction:** Our IV needs to only move Y through eligibility

Instrument of choice: simulated instrument

- Share of kids that would be eligible for Medicaid in state s in year t
- The first stage should be positive

Currie and Gruber (1996): Estimation approach

With the instrument, we simply estimate:

$$D_{ist} = \theta Z_{ist} + \alpha_s + \delta_t + \varepsilon_{ist}$$

$$Y_{ist} = \tau \hat{D}_{ist} + \alpha_{id} + \delta_t + \varepsilon_{ist}$$

Currie and Gruber (1996): Estimation approach

With the instrument, we simply estimate:

$$D_{ist} = \theta Z_{ist} + \alpha_s + \delta_t + \varepsilon_{ist}$$

$$Y_{ist} = \tau \hat{D}_{ist} + \alpha_{id} + \delta_t + \varepsilon_{ist}$$

Identifying assumption: Conditional on fixed effects, simulated eligibility doesn't affect health other than through actual eligiblity

- Is this reasonable?

Currie and Gruber (1996): Estimation approach

With the instrument, we simply estimate:

$$D_{ist} = \theta Z_{ist} + \alpha_s + \delta_t + \varepsilon_{ist}$$

$$Y_{ist} = \tau \hat{D}_{ist} + \alpha_{id} + \delta_t + \varepsilon_{ist}$$

Identifying assumption: Conditional on fixed effects, simulated eligibility doesn't affect health other than through actual eligiblity

- Is this reasonable?

LATEs: Estimates are the LATE for whom?

- How should this compare to ATE?

Currie and Gruber (1996): Utilization

TABLE IV
MEDICAID ELIGIBILITY AND THE UTILIZATION OF MEDICAL CARE
LINEAR PROBABILITY MODELS: COEFFICIENTS × 10^2

Dependent var	(1) OLS No visit last year	(2) OLS Visit last 2 weeks	(3) OLS Hospital last year	(4) TSLS No visit last year	(5) TSLS Visit last 2 weeks	(6) TSLS Hospital last year
Medicaid	−2.510	−0.119	0.681	−9.553	4.853	3.960
eligibility	(0.309)	(0.237)	(0.153)	(3.037)	(2.803)	(1.646)
Male	−0.034	0.691	0.763	−0.033	0.691	0.763
	(0.159)	(0.132)	(0.078)	(0.159)	(0.132)	(0.078)
Black	4.149	−3.354	−0.611	4.362	−3.505	−0.710
	(0.260)	(0.195)	(0.123)	(0.276)	(0.212)	(0.133)
Hispanic	1.738	−0.922	0.019	1.978	−1.093	−0.093
	(0.294)	(0.234)	(0.140)	(0.311)	(0.254)	(0.150)
Mom is HS	2.809	−0.613	0.264	3.255	−0.927	0.057
dropout	(0.246)	(0.180)	(0.118)	(0.316)	(0.252)	(0.157)
Mom has some	−3.098	1.177	−0.263	−3.269	1.298	−0.183
college	(0.197)	(0.175)	(0.098)	(0.210)	(0.188)	(0.064)
Dad is HS	3.069	−0.832	−0.216	3.365	−1.041	−0.354
dropout	(0.296)	(0.212)	(0.137)	(0.323)	(0.243)	(0.154)
Dad has some	−2.392	0.672	−0.252	−2.378	0.662	−0.258
college	(0.223)	(0.191)	(0.111)	(0.223)	(0.192)	(0.109)
Child is oldest	−2.540	0.990	−0.049	−2.372	0.872	−0.127
	(0.197)	(0.157)	(0.092)	(0.210)	(0.171)	(0.099)
Number of	1.610	−0.640	−0.234	2.111	−0.936	−0.430
siblings	(0.095)	(0.066)	(0.040)	(0.204)	(0.175)	(0.105)
No male head	−5.243	2.195	0.618	−4.985	2.012	0.498
	(0.395)	(0.315)	(0.196)	(0.410)	(0.332)	(0.204)
Mom is	−0.214	1.434	−0.445	0.027	1.264	−0.556
respondent	(0.569)	(0.541)	(0.349)	(0.579)	(0.549)	(0.352)

Currie and Gruber (1996): Site of care

TABLE V
MEDICAID ELIGIBILITY AND THE SITE OF CARE
ALL REGRESSIONS RUN AS INSTRUMENTAL VARIABLES
MEDICAID ELIGIBILITY COEFFICIENT AND MEANS ARE × 100

	(1) Doctor's office	(2) ER or hospital outpatient clinic	(3) Other site
Medicaid eligibility	5.073	1.174	−1.217
	(2.479)	(1.117)	(1.100)
Mean of dependent var	8.707	1.666	1.473
Number of obs.	227169	227169	227169

Standard errors are in parentheses. All regressions also include all the variables listed in Table V, as well as an intercept; dummy variables for each state, calendar year, and year of age; season dummies; interactions between calendar year and year of age dummies; and interactions between year of age and state dummies. Eligibility is instrumented using simulated eligibility calculated from the CPS, and matched to individuals by state, year, and age. Standard errors are corrected for heteroskedasticity.

Currie and Gruber (1996): Health

TABLE VI
EFFECTS OF MEDICAID ELIGIBILITY ON CHILD MORTALITY
DEPENDENT VARIABLE IS DEATHS PER 10,000 CHILDREN

	(1) All causes	(2) Internal causes	(3) External causes
Percent eligible	−1.277	−1.016	−0.261
	(0.482)	(0.359)	(0.363)
Mean of dep var	3.807	1.926	1.881
Number of obs	816	816	816

Standard errors are in parenthese. Dependent variable is death rate per 10,000 children in state/year/race/age group, where age groups are 1–4 years old and 5–14 years old. Regressions are run as instrumental variables, where percent eligible in state/year/age group cell is instrumented using simulated eligibility in that cell. Regressions include state, year, and age group dummies. Standard errors are corrected for heteroskedasticity.

Hanratty (1996): Canadian DD version

Table 1—Selected Characteristics of Provinces in 1961 by Date of Implementation of Universal Health Insurance Program

Province	Date adopted program	Infant deaths/ live births	Low-birth-weight births/live births	Family income below $3,000	Post high school (age 20 or more)	Urban	GDP per capita/Canada average GDP per capita
Saskatchewan	July 1962	0.026	0.068	0.307	0.068	0.430	0.780
British Columbia	July 1968	0.024	0.068	0.195	0.152	0.726	1.114
Nova Scotia	April 1969	0.028	0.069	0.355	0.068	0.543	0.653
Newfoundland	April 1969	0.038	NA	0.494	0.045	0.507	0.502
Manitoba	April 1969	0.025	0.067	0.228	0.096	0.639	0.901
Alberta	July 1969	0.027	0.084	0.196	0.119	0.633	1.087
Ontario	October 1969	0.023	0.076	0.167	0.142	0.773	1.202
Quebec	November 1970	0.031	0.089	0.208	0.120	0.743	0.906
Prince Edward Island	December 1970	0.033	0.056	0.427	0.069	0.324	0.494
New Brunswick	January 1971	0.026	0.065	0.372	0.068	0.465	0.602

Notes: Yukon and Northwest Territories introduced programs in April 1972 and April 1971, respectively. In addition, Alberta (October 1963), Ontario (July 1966) and British Columbia (September 1965) implemented voluntary health insurance subsidy progams prior to universal insurance. For more information, see Andreopoulos (1975), Shillington (1972) and Department of National Health and Welfare (1966). Dollars are reported in 1961 Canadian dollars (3,000 Canadian 1961 dollars equals 12,600 1993 U.S. dollars).

Hanratty (1996): Canadian DD version

TABLE 3—GROUPED LOGIT ESTIMATES OF IMPACT OF NATIONAL HEALTH INSURANCE ON INFANT MORTALITY RATE

Variable	Mean	Model 1	Model 2	Model 3
NHI	0.368	−0.046	−0.040	−0.032
	(0.353)	(0.014)	(0.021)	(0.031)
NHI last year	0.311			−0.003
	(0.338)			(0.027)
NHI in one year	0.426			−0.012
	(0.363)			(0.026)
Age 15–19	0.211	2.216	2.848	2.844
	(0.028)	(0.545)	(0.565)	(0.566)
Age 35–44	0.285	−2.835	−3.079	−3.088
	(0.024)	(0.427)	(0.429)	(0.429)
Income < \$3,000	0.236	0.435	0.636	0.637
	(0.082)	(0.172)	(0.177)	(0.177)
English or French	0.735	0.738	0.998	0.999
	(0.135)	(0.267)	(0.274)	(0.274)
Urban	0.729	−0.338	−0.181	−0.183
	(0.179)	(0.120)	(0.126)	(0.126)
Log real provincial GDP per capita	9.202	−0.016	0.103	0.109
	(0.219)	(0.070)	(0.091)	(0.093)
Post high school	0.136	0.136	−0.637	−0.644
	(0.039)	(0.201)	(0.268)	(0.272)
Trend	6.958	−0.056		
	(3.520)	(0.004)		
Infant mortality rate	0.022			
	(0.006)			
Adjusted R^2		0.760	0.764	0.764
County effects		Yes	Yes	Yes
Year effects		No	Yes	Yes

Hanratty (1996): Canadian DD version

TABLE 4—LOGIT ESTIMATES OF THE IMPACT NATIONAL HEALTH INSURANCE ON LOW BIRTH WEIGHT

	Mean	Model 1	Model 2	Model 3	Model 4
NHI	0.329	−0.032	−0.015	−0.022	0.001
		(0.008)	(0.010)	(0.015)	(0.029)
NHI last year	0.267			0.021	
				(0.013)	
NHI next year	0.391			−0.010	
				(0.013)	
NHI × single parent	0.030				−0.112
					(0.031)
Single parent	0.070	0.414	0.414	0.414	0.263
		(0.012)	(0.012)	(0.012)	(0.091)
Log real provincial GDP per capita	9.199	−0.011	−0.043	−0.031	−0.047
		(0.042)	(0.049)	(0.050)	(0.049)
Single birth	0.980	−2.897	−2.897	−2.897	−2.896
		(0.006)	(0.006)	(0.006)	(0.006)
Male	0.514	−0.200	−0.200	−0.200	−0.200
		(0.003)	(0.003)	(0.003)	(0.003)
Father's age < 20	0.015	0.214	0.214	0.214	0.211
		(0.015)	(0.015)	(0.015)	(0.015)
Father's age 20–29	0.491	−0.007	−0.007	−0.007	−0.007
		(0.008)	(0.008)	(0.008)	(0.008)
Father's age 30–39	0.346	−0.080	−0.079	−0.079	−0.079
		(0.007)	(0.007)	(0.007)	(0.007)
Father's age missing	0.059	0.031	0.030	0.030	0.049
		(0.014)	(0.014)	(0.014)	(0.015)
Mother's age < 20	0.104	−0.022	−0.022	−0.022	−0.019
		(0.013)	(0.013)	(0.013)	(0.013)
Mother's age 20–29	0.616	−0.178	−0.178	−0.178	−0.178
		(0.011)	(0.011)	(0.011)	(0.011)
Mother's age 30–39	0.252	−0.104	−0.104	−0.104	−0.105
		(0.011)	(0.011)	(0.011)	(0.011)
1st live birth	0.337	0.232	0.231	0.232	0.228
		(0.006)	(0.006)	(0.006)	(0.006)
2nd live birth	0.264	0.094	0.094	0.094	0.093
		(0.005)	(0.005)	(0.005)	(0.005)
3rd live birth	0.162	0.018	0.018	0.018	0.017
		(0.006)	(0.006)	(0.006)	(0.006)
Past stillbirths	0.031	0.752	0.753	0.753	0.753
		(0.008)	(0.008)	(0.008)	(0.008)
Low birth weight	0.073				
Log-likelihood		−1424259	−1424220	−1424218	−1424129

Recap

TL;DR:

1. Currie and Gruber (1996) is a seminal study of the effects of Medicaid expansions
2. Finds that Medicaid dramatically increases utilization and health
3. Uses an IV strategy based on simulated instruments (credible?)
4. Hanratty: positive impacts of insurance on health in Canada