

# Health, Human Capital & Education

- 1 Impact of Education on Health
- 2 Human Capital Effect of Hepatitis B Vaccination
- 3 Effect of Air Pollution on Health

Kamhon Kan/簡錦漢  
*Academia Sinica*/中研院 經濟所

科技部人社中心 學術研習營  
國立中正大學 經濟學系

2019.06.06

## *Timetable*

- 10:00-11:00 Lecture
- 11:00-11:20 Discussion/Recess
- 11:20-12:20 Lecture
- 12:20-14:00 Lunch/Discussion
- 14:00-15:00 Lecture
- 15:00-16:00 Discussion

# Environment, Public Policy and Human Capital

## The Impact of Education on Health

## *Emphasis*/重點

- Effect of education
- Causal inference - instrumental variables/Regression Discontinuity
- NHI database (健保資料庫)

# *Motivation*

- Social scientists have long been interested in the education-health relationship
  - Education policies: Health benefits need to be considered in education-related programs?
  - Public Health policies: Education is a power tool to improve health?

## *Possible channels of the effect*

- Possible channels for education to affect health
  - **Knowledge:** Enhances an individual's health production efficiency and allocative efficiency (see Grossman, 2004, for a review and exposition).
  - **Attitude:** Reduces one's rate of time preference—→more future-oriented and more likely to invest in health (e.g., Becker and Mulligan, 1997)
  - **Income:** Have better access to higher quality inputs of health production
  - **Occupations:** Not having jobs which are harmful to health (less risky or stressful).

## Possible endogeneity

- Positive associations are found in literature (see a review by Grossman, 2006)

$$\text{health}_i = \text{schooling}_i \underset{(+)}{\theta} + \mathbf{X}_i \boldsymbol{\delta} + \text{unobservables}_i$$

- May not be causal—endogeneity
  - **Omitted variables:** Affect both (e.g., time preferences, ability)
  - **Reverse causality:**
    - ▶ Healthier individuals → obtain more education.  
Healthier → efficient learners.
    - ▶ Healthier → more incentive to invest in education.  
Longer life expectancies → enjoy greater payoffs.
- Endogeneity needs to be accounted for to obtain a causal effect.

## *Recent studies — Natural experiments*

- Recent studies: use quasi-experiments for identification
- E.g., Lleras-Muney (2005), Arendt (2005), Oreopoulos (2006),
- Results are mixed → more research needed.
- Current study: better data, comprehensive analysis, and more effective experiment.



# *What Has Been Done*

## ■ Earlier studies

- Effect of education on mortality, self-reported health, smoking, obesity, etc.
- IV or natural experiments for identification; survey data.

## ■ Recent studies

- Focus on mortality:  
Clark and Royer (2010) - the UK  
Meghir et al. (2018) - Sweden  
Mazumder (2008) and Black et al. (2015) - U.S.
- Using population data; natural experiments for identification.
- Find no effects. Maybe because of low return to education (Galama *et al.*, 2018)

# *What Has Been Done*

- **Mortality**

⇒ Albouya and Lequien (2009), Lleras-Muney (2005), van Kippersluis *et al.* (2011), and Clark and Royer (2013)

- **Self-reported health status or diseases**

⇒ Arendt (2005), Silles (2009), Kemptner *et al.* (2011), and Clark and Royer (2013)

- **Health behavior—Smoking**

⇒ Grimard and Parent (2007), Tenn *et al.* (2010) and Webbinka *et al.* (2010)

- **Indirect measure of health—Weight & BMI**

⇒ Webbinka *et al.* (2010) and Kemptner *et al.* (2011)

## *Room for Writing Another Paper?*

## *Room for Writing Another Paper?*

- Precise & direct measurement of health conditions (self-reported/indirect measures in previous studies) and mortality.
  - Use administrative records/physicians' diagnosis and death registry.
  - Study mechanisms — income, occupation, health behavior, and medical services utilization?
- Recent studies find no effects, all are in contexts with low return to education. See Galama *et al.* (2018).

# Health outcomes examined

- **Mortality**
- **Cardiovascular diseases:**
  - ⇒ high blood pressure, heart disease, stroke, etc. (risk factors: obesity, dietary habit, exercise)
- **Other chronic diseases:**
  - ⇒ type 2 diabetes, type 1 diabetes (falsification) liver diseases, lung diseases, high blood pressure (risk factors: smoking, alcohol, dietary habit)
- **Smoking-related diseases:**
  - ⇒ Any of these diseases: coronary heart disease, cancer (Laryngeal/喉癌, lung cancer/肺癌, Pleural cancer/胸膜癌), cerebrovascular disease (腦血管疾病, e.g., stroke), chronic obstructive pulmonary disease (慢性阻塞性肺病)
- **Others:**
  - ⇒ injury, occupation diseases, cancer, accident (risk factors: risk attitude, occupation)
- **Utilization:**
  - ⇒ inpatient, outpatient, preventive care (health examination) (hints on severity and why/how education affects health)

## *Taiwan's compulsory education*

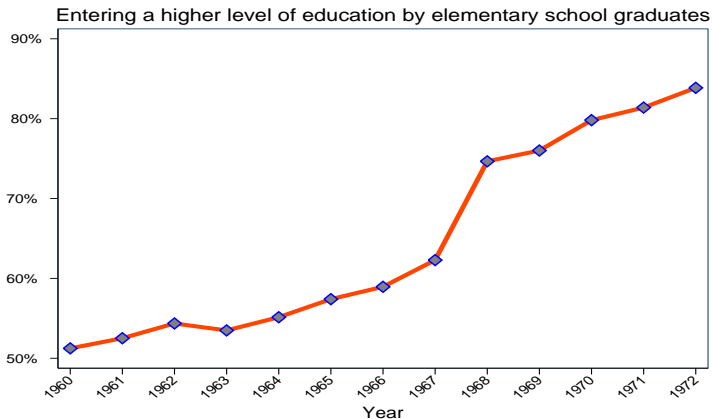
- Compulsory education since 1943 (colony of Japan)
  - children 6–12 received six-year free education
- 9 year compulsory education since 1968 (children under 15)
  - To sustain economic growth and upgrade industrial structure
- Part of 4<sup>th</sup> Economic Development Plan (1965–68) of Taiwan's Council for Economic Development and Planning
  - construction of new schools (243 new junior high schools opened 1966–1968)
  - increasing the number of qualified teachers (teachers per thousand of elementary school graduates increased by 31.2% 1966–1968.)

## *Effectiveness of the Reform*

- Mid 1960s, 60% elementary school graduates continued education
- Junior high school enrollment rate jumped, 62.29% → 74.66% (1967 and 1968).
- Enrollment rate reached 83.86% in 1972

# Effectiveness of the Reform

Effect on entering a higher level of education





# Data

## 2000 Population Census and Death Records

- Entire population:
  - 2000 Population Census (education)
  - + 2001–2015 Death Registry (mortality)
  - + National Health Insurance Data (medical records, income, occupation, ...)
  - Linked by ID
  - Exclude non-citizens
- Have exact date-of-birth
- Look at individuals born surrounding September 2, 1955.
  - ▲ Born before 9/2/1955: 13 or older in Sept 1968. Not affected.
  - ▲ Born after 9/2/1955: 12 or younger in Sept 1968. Affected by reform.

# Data

## Health outcomes

- Mortality: Whether an individual died during 2001-2015
  - ⇒ Records in death registry.
- Diseases (e.g., diabetes, heart diseases, cardiovascular diseases, smoking-related diseases, injury, dementia).
  - ⇒ Major diagnosis (ICD 9 codes) for each inpatient/outpatient visit.
  - ⇒ 3 outpatient visits or 1 inpatient visit within 90 days → have a disease.
  - ⇒ Other criteria experimented (e.g., 4 outpatient visits in 180 days).

# Descriptive Statistics

## 2001–2015 mortality

Table: Individuals aged 12 and 13 on September 3, 1968<sup>†</sup>

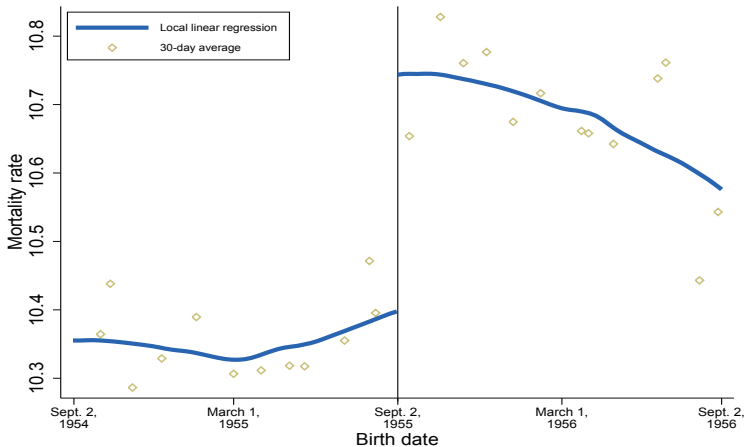
	Men		Women	
	Aged <b>12</b> on Sep. 2, 1968 (Affected)	Aged <b>13</b> (Not affected)	Aged <b>12</b> on Sep. 2, 1968 (Affected)	Aged <b>13</b> (Not affected)
Mortality rate (Death Registry)	0.0960 ( 0.3033)	0.1025 ( 0.2947)	0.0403 ( 0.2040)	0.0435 ( 0.1966)
Schooling years (2000 Census)	10.6850 ( 3.5920)	10.3544 ( 3.4148)	9.3739 ( 3.6561)	8.9846 ( 3.6049)
Observations	181718	176739	178062	172904

<sup>†</sup> Standard deviations in parentheses and number of observations in square brackets.

# Preliminary exploration

## Men's schooling years

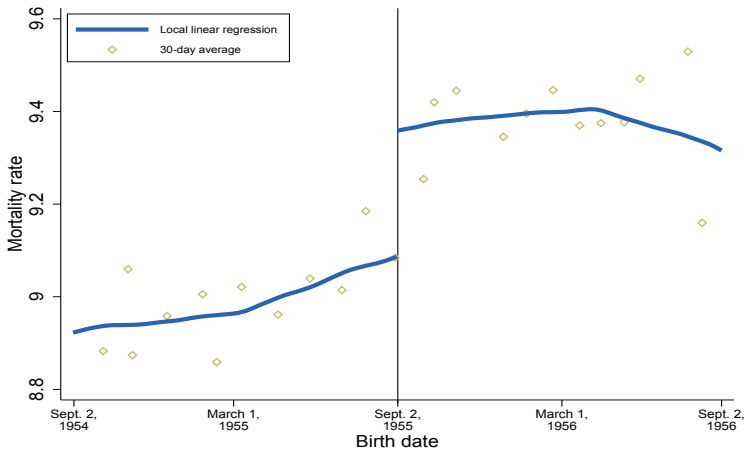
Figure: Years of schooling of cohorts born 1954–1956 — Men



# Preliminary exploration

## Women's schooling years

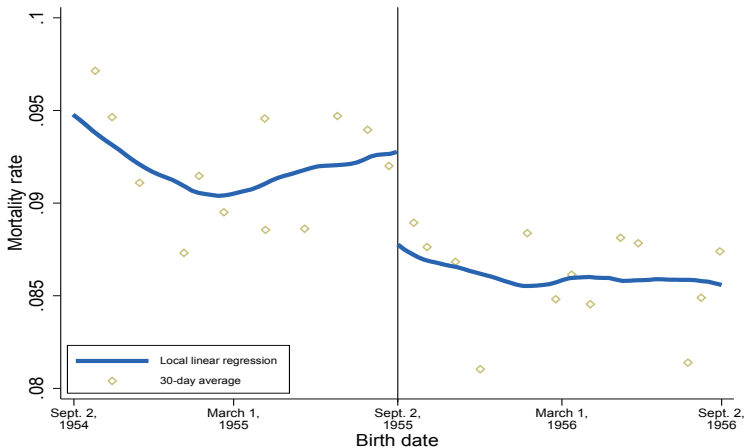
Figure: Years of schooling of cohorts born 1954–1956 — Women



# Preliminary exploration

## Men's mortality

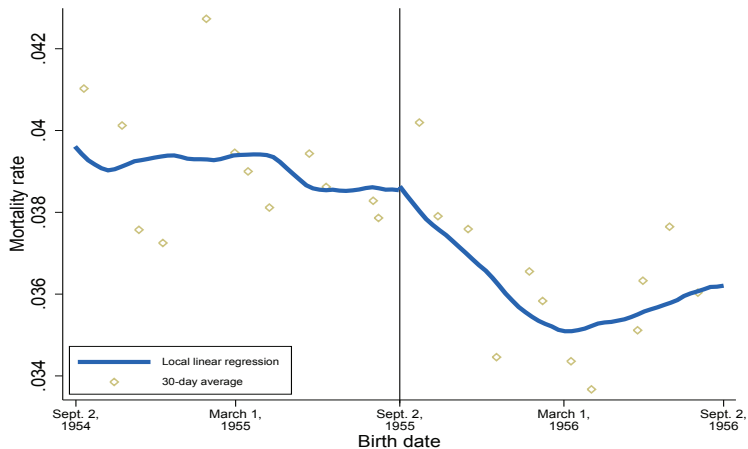
Figure: Mortality rate — Men



# Preliminary exploration

## Women's mortality

Figure: Mortality rate — Women



# Empirical Strategy

## Relationship of interest

### ■ Model of interest

$$d_i = \delta + s_i\theta + \mathbf{Z}_i\boldsymbol{\delta} + \epsilon_{it}, \quad (1)$$

- $d_i$  = died in 2001–2015 or having a disease
  - $s_i$  = schooling years
  - $\mathbf{Z}_i$  = control variables (e.g., county dummies, age, ...)
  - $\epsilon_i$  = error term
- Omitted variables or reverse causality:  $\text{Corr}(s_i, \epsilon_i) \neq 0$ , i.e.,  $s_i$  is endogenous.



# Empirical Strategy

## Regression Discontinuity—Fuzzy Design

$$\theta = \frac{\lim_{e \uparrow 0} E(d_{it} | R_i = e) - \lim_{e \downarrow 0} E(d_{it} | R_i = e)}{\lim_{e \uparrow 0} E(s_{it} | R_i = e) - \lim_{e \downarrow 0} E(s_{it} | R_i = e)}, \quad (2)$$

■ Estimated by 2SLS  $d_i = \kappa_0 + s_i\theta + f(R_i) + \mathbf{X}_i\boldsymbol{\lambda} + u_i$  (3)

$$s_i = \gamma_0 + D_i\beta + q(R_i) + \mathbf{X}_i\boldsymbol{\tau} + v_i \quad (4)$$

- $R_i$  = days to reach 13 in 9/1968 (i.e., running variable)
- $D_i = 1$ : younger than 13 in 9/1968 (i.e., affected by law change), instrument
- $f(\cdot)$ ,  $q(\cdot)$  = continuous function of  $R_i$
- $\mathbf{X}_i$ : predetermined variables (county dummies)
- Local linear regression (**bandwidth**: Imbens & Kalyanaraman, 2011; **kernel**: triangular); **SE** birth-date clustered via 2SLS

# Main Results

## 2001–2015 mortality

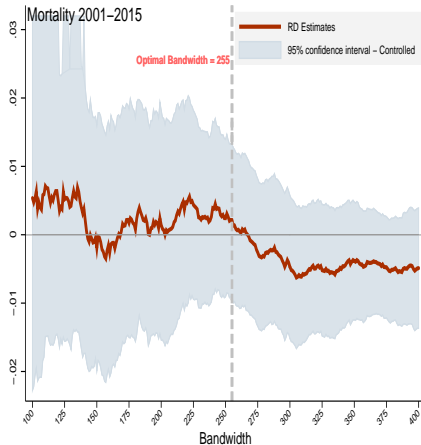
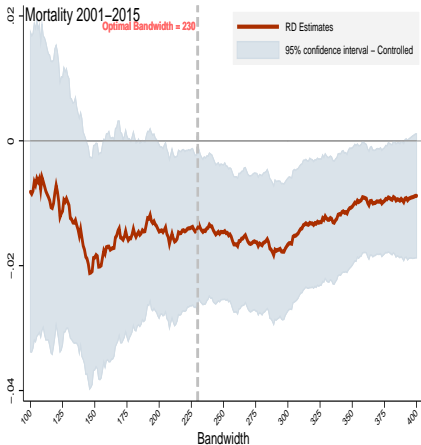
Table: Regression Discontinuity Design  
Estimates<sup>†</sup>

	Men	Women
Bandwidth <sup>‡</sup>	293 days	325 days
Effect of schooling years on mortality (RD)	-0.0152** ( 0.0062)	-0.0009 (0.0059)
Effect of law change on mortality	-0.0055** ( 0.0024)	-0.0002*** (0.0016)
Effect of law change on schooling years	0.3646*** ( 0.0511)	0.2644*** (0.0474)
<i>F</i> -Statistic of law change	50.8455	31.0748
Effect of schooling years on mortality (OLS)	-0.0103*** (0.0002)	-0.0025*** (0.0001)
Mean mortality rate	0.0980	0.0416
Observations	278912	302153

# Main Results

## 2001–2015 mortality

Figure: Effect of schooling on mortality



# *Mechanism*

## The Role of Income

# Mechanism

## Controlling for Income & Occupation

### ■ Controlling for Income & Occupation

$$d_i = \delta_0 + s_i\theta + R_i\delta_1 + D_i \times R_i\delta_2 + \mathbf{X}_i\boldsymbol{\delta}_3 + \sum_3^1 occ_{ji}\delta_{3j} + \sum_3^1 income_{ji}\delta_{4j} + error_i$$

- $income_{1i}$  — year 2000 income below the bottom 25th percentile
- $income_{2i}$  — year 2000 income above the 75th percentile
- $income_{3i}$  — year 2000, no income
- $occ_{1i}$  — public sector
- $occ_{2i}$  — private sector
- $occ_{3i}$  — employer/self-employed
- $occ_{4i}$  — dependent

# Mechanism

2001–2015 mortality: Effect of Income and Occupation<sup>†</sup>

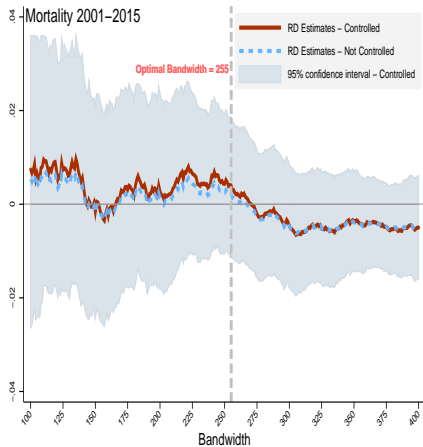
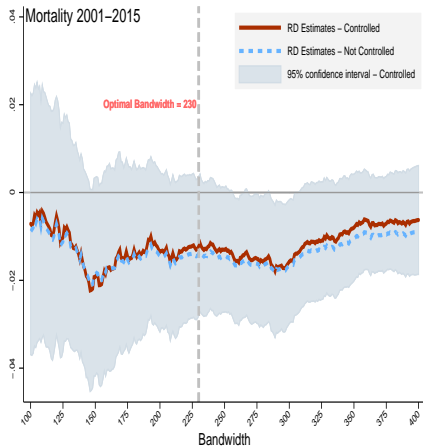
	Men			Women		
	Not		Controlled	Not		Controlled
	Controlled	Controlled		Controlled	Controlled	
	Bandwidth: 293 days			Bandwidth: 325 days		
	[1]	[2]	[3]	[4]	[5]	[6]
Effect of Education (RD)	-0.0152** (0.0062)	-0.0149** (0.0072)	-0.0140* (0.0081)	0.0021 (0.0057)	-0.0003 (0.0070)	0.0001 (0.0075)
Public Sector		-0.0377 (0.0291)	-0.0206* (0.0209)		-0.0264*** (0.0262)	-0.0184* (0.0196)
Private Sector		-0.0369*** (0.0051)	-0.0330*** (0.0041)		-0.0130* (0.0067)	-0.0117** (0.0047)
Employer/Self-Employed		-0.0219* (0.0121)	-0.0047** (0.0025)		-0.0129* (0.0137)	-0.0027* (0.0026)
Dependent		-0.0460*** (0.0169)	-0.0482*** (0.0180)		-0.0176*** (0.0211)	-0.0205*** (0.0216)
Income below 25th percentile			0.0254*** (0.0094)			0.0062*** (0.0074)
Income above 75th percentile			-0.0044*** (0.0145)			-0.0107* (0.0200)
No income		0.0279* (0.0149)	0.0494** (0.0203)		0.0130** (0.0215)	0.0173** (0.0204)
Not working & not dependent		—	—		—	—
Observations		219451			237439	

<sup>†</sup> Effect of education reduced after controlling for income & occupation (-0.0145 → -0.0135).

# Main Results—Income & occupation controlled

## Reduced Form Estimation

Figure: Income & occupation controlled



## Estimation by IV

### Program intensity as IVs

#### ■ 2SLS

$$d_{ic} = \delta_0 + s_{ic}\delta_1 R_i \delta_2 + D_i R_i \delta_3 + \mathbf{X}_{ic} \boldsymbol{\lambda}_4 + \text{error}_i \quad (5)$$

$$s_{ic} = \lambda_0 + P_{ic} \lambda_1 + R_i \lambda_2 + D_i R_i \lambda_3 + \mathbf{H}_{ic} \boldsymbol{\lambda}_4 + \text{error}_i \quad (6)$$

- $H_{ic}$  number of schools in county  $c$  when individual  $i$  was aged 12, and county dummies
- $X_{ic}$  county dummies
- $c$  county where  $i$  was born.

#### ■ Results:

	Men	Women
Effect of education	0.0001 (0.0006)	-0.0005 (0.004)
F-Statistic	136.7918	183.0845
Observations	1102290	1091901



## Discrete running variable

Month-of-birth

BW (months)	Effect of Schooling	
	Men (optimal BW=8)	Women (optimal BW=10)
6	-0.0289	-0.0171
7	-0.0464*	-0.0179
8	-0.0408**	-0.0135 (0.0083)
9	-0.0446** (-0.0136)	-0.0121
10	-0.0489**	-0.0114
11	-0.0506**	-0.0130
12	-0.0498**	-0.0159
13	-0.0636	-0.0258
14	-0.1013	-0.0486
15	-0.1251	-0.1010
16	-0.1311	-0.0774
17	-0.1297	-0.0904
18	-0.1163	-0.0925
19	-0.0989	-0.0431
20	-0.0847	-0.0221
21	-0.0777	-0.0176
22	-0.0739	-0.0184
23	-0.0688	-0.0195
24	-0.0524*	-0.0118

## Main Results—Diseases

# Main Results—Diseases

## Chronic Diseases

Table: Chronic Diseases

	Diabetes, type 1		Diabetes, type 2		Congenital diseases		Liver diseases		Lung diseases		Cancer	
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]
	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women
BW	332	394	386	364	249	257	252	319	334	335	325	278
RD	0.0004 ( 0.0008)	-0.0006 ( 0.0010)	-0.0026 ( 0.0058)	-0.0183** ( 0.0080)	0.0024 ( 0.0019)	-0.0026 ( 0.0034)	-0.0051** ( 0.0022)	-0.0030 ( 0.0018)	0.0038 ( 0.0046)	-0.0030 ( 0.0053)	-0.0022 ( 0.0060)	0.0139 ( 0.0087)
OLS	-0.0001*** ( 0.0000)	-0.0001*** ( 0.0000)	-0.0029*** ( 0.0002)	-0.0072*** ( 0.0002)	-0.0001* ( 0.0000)	-0.0005*** ( 0.0001)	-0.0007*** ( 0.0001)	0.0002*** ( 0.0000)	-0.0035*** ( 0.0001)	0.0020*** ( 0.0001)	-0.0037*** ( 0.0001)	0.0007*** ( 0.0002)
Mean	0.0015 ( 0.0383)	0.0013 ( 0.0356)	0.1073 ( 0.3095)	0.0928 ( 0.2901)	0.0049 ( 0.0701)	0.0084 ( 0.0910)	0.0074 ( 0.0858)	0.0027 ( 0.0523)	0.0384 ( 0.1921)	0.0271 ( 0.1623)	0.0778 ( 0.2679)	0.0770 ( 0.2665)

# Main Results—Diseases

## Cardiovascular Diseases

Table: Cardiovascular Diseases

	Cardiovascular diseases		Heart diseases		Stroke		AMI		High Blood Pressure	
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women
BW	254	331	285	295	254	297	225	197	206	227
RD	-0.0002 ( 0.0116)	-0.0084 ( 0.0128)	0.0005 ( 0.0053)	-0.0118* ( 0.0061)	-0.0032 ( 0.0050)	0.0066 ( 0.0049)	-0.0032 ( 0.0041)	-0.0025 ( 0.0036)	0.0026 ( 0.0124)	-0.0058 ( 0.0132)
OLS	-0.0001 ( 0.0003)	-0.0095*** ( 0.0002)	-0.0008*** ( 0.0001)	-0.0016*** ( 0.0001)	-0.0025*** ( 0.0001)	-0.0016*** ( 0.0001)	-0.0005*** ( 0.0001)	-0.0005*** ( 0.0001)	0.0028*** ( 0.0003)	-0.0073*** ( 0.0002)
Mean	0.2829 ( 0.4504)	0.2358 ( 0.4245)	0.0536 ( 0.2251)	0.0328 ( 0.1782)	0.0395 ( 0.1947)	0.0209 ( 0.1431)	0.0219 ( 0.1463)	0.0053 ( 0.0728)	0.1674 ( 0.3733)	0.1656 ( 0.3717)

# Main Results—Diseases

## Occupational Accidents & Injury

Table: Occupational Accidents and Injuries

	Occupational accident		Injuries	
	[1]	[2]	[3]	[4]
	Men	Women	Men	Women
Optimal Bandwidth	315	299	299	231
Effect of Education	-0.0046** ( 0.0024)	0.0006 ( 0.0026)	-0.0024 ( 0.0083)	-0.0174 ( 0.0161)
Effect of Education (OLS)	-0.0005*** ( 0.0001)	-0.0001* ( 0.0000)	-0.0132*** ( 0.0002)	-0.0087*** ( 0.0002)
Mean of Dependent Variable	0.0097 ( 0.0978)	0.0073 ( 0.0852)	0.1815 ( 0.3854)	0.1649 ( 0.3711)

# Main Results—Diseases

## Medical Care Utilization

Table: Medical Care Utilization

	# Outpatient Visits <sup>‡</sup> (in log)		Medical Utilization - Screening/Tests (#points <sup>‡</sup> , in log)		Medical utilization - Medication (#points <sup>‡</sup> , in log)		# Inpatient visits <sup>‡</sup> (in log)		# Days Hospitalized <sup>‡</sup> (in log)		Inpatient Services <sup>‡</sup> - Total (#points <sup>ddag</sup> , in log)		# Physical Examinations <sup>‡</sup> (#times, in log)	
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]	[14]
	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women
<b>BW</b>	347	236	320	199	257	323	312	170	203	261	189	174	326	298
<b>RD</b>	0.0106 ( 0.0121)	-0.0025 ( 0.0243)	0.1248* ( 0.0679)	-0.0362 ( 0.1452)	0.0455 ( 0.0841)	-0.0238 ( 0.0937)	-0.0371*** ( 0.0138)	0.0109 ( 0.0291)	-0.0444*** ( 0.0157)	0.0155 ( 0.0165)	-0.3057* ( 0.1818)	0.2078 ( 0.2725)	-0.0010 ( 0.0047)	0.0089 ( 0.0112)
<b>OLS</b>	-0.0027*** ( 0.0003)	-0.0128*** ( 0.0004)	0.0012 ( 0.0018)	-0.0685*** ( 0.0025)	0.0064*** ( 0.0020)	-0.0664*** ( 0.0019)	-0.0220*** ( 0.0004)	-0.0082*** ( 0.0005)	-0.0042*** ( 0.0003)	-0.0019*** ( 0.0003)	-0.1908*** ( 0.0042)	-0.1045*** ( 0.0042)	0.0007*** ( 0.0001)	0.0034*** ( 0.0002)
<b>Mean</b>	0.3523 ( 0.5820)	0.4301 ( 0.6249)	2.8178 ( 3.1584)	3.3300 ( 3.1818)	2.5879 ( 3.1930)	2.8071 ( 3.1108)	0.2575 ( 0.6463)	0.1756 ( 0.5301)	0.0550 ( 0.4724)	0.0350 ( 0.3766)	4.7369 ( 5.4806)	4.1763 ( 5.2660)	0.0492 ( 0.2163)	0.1294 ( 0.3357)

# Main Results—Diseases

## Dental Care

Table: Dental Diseases and Service Utilization

	Scaling (# visits <sup>‡</sup> , in log)		Periodontosis		Caries	
	[1]	[2]	[3]	[4]	[5]	[6]
	Men	Women	Men	Women	Men	Women
BW	284	314	153	311	259	266
RD	0.0125 ( 0.0082)	0.0095 ( 0.0108)	0.0185 ( 0.0171)	0.0045 ( 0.0136)	0.0094 ( 0.0081)	0.0109 ( 0.0156)
OLS	0.0082*** ( 0.0002)	0.0066*** ( 0.0002)	0.0171*** ( 0.0004)	0.0132*** ( 0.0003)	0.0078*** ( 0.0002)	0.0057*** ( 0.0003)
Mean	0.1109 ( 0.3140)	0.1295 ( 0.3358)	0.2744 ( 0.4462)	0.2868 ( 0.4523)	0.1537 ( 0.3607)	0.2090 ( 0.4066)

## *Conclusion and interpretation*

- Education has a large impact on men's mortality, but not on women's.
  - ⇒ Women is in a better health condition than men for given age.
- Some of this effect is via education's impact on income & occupation.
- Education reduces the probability of having some adverse health conditions (e.g., occupation accidents, heart, diabetes, liver).
  - Occupation & behavior matters
- Better education individuals does not consumes more outpatient services. They consume less inpatient services.
  - ⇒ Better educated individuals are less likely to have serious conditions (hospitalization).
- Better educated individuals not necessarily use more preventive care.