Does retirement make people happy? A mental health perspective

Kai Chiu Yang / Tinghuan Liu

Program Evaluation and Causal Inference

2021/4/27

CONTENTS

1. Policy Background

Statutory Retirement Age

Individuals participating in **pension programs** and reaching **the statutory age** can apply for retirement.

Age	Males	Females
Statutory Retirement	60	50
Officials in the gov. & Inst.	60	55
Workers doing hazardous jobs*	55	45
Disabled to work**	50	45

Retiring at the statutory retirement age can be regarded as compulsory in practice.

- Retirees only receive monthly pension **AFTER** their application are approved.
- The monthly pension between the date of age 60 (for men) and the date of retired will **NOT** be repaid.
- firms are **NOT allowed to employ** individuals once they reach the statutory retirement age.
- Those who are willing to work after retirement can choose to provide dependent services to firms, instead of signing a labor contract, but it could be risky in law and thus relatively rare.

*E.g., working at heights, in high temperatures, etc.
** Needs to be certificated by medical department

2. Empirical Design

Methodology: Fuzzy RD

- **Three possible cutoffs** in terms of retirement status: age 60, 55 and 50 for men, and age 55, 50 and 45 for women.
- **Not everyone gets retired exactly at the date of the cutoff age**, as one might retire early, be re-employed without signing a labor contract after retirement, or just simply forget to apply for retirement.

$$\underbrace{depre_i}_{\text{A measure of depression level}} = \alpha_0 + \alpha_1 \boxed{D_i} + \alpha_2 \underbrace{(age_i - c)}_{\text{Age Effect wrt. diff. cutoffs}} + \alpha_3 (age_i - c)\boxed{Z_i} + f\boxed{(income_i)} + \underbrace{u_i}_{\text{Clustered at individual level}}$$

D_i	A dummy equal to 1 if individual i is retired
Z_i	A dummy (instrument) equal to 1 if individual i's age is above the cutoff
$f(income_i)$	An RD polynomial in individual i's income

- **Data:** *China Health and Retirement Longitudinal Study* (CHARLS) of the year 2013, 2015 and 2018.
- **Baseline:** linear specification and a bandwidth of 1 year.
- We combine data of year 2013, 2015 and 2018, and estimate coefficients in pooling OLS.

2. Empirical Design

Methodology: Fuzzy RD

- **Three possible cutoffs** in terms of retirement status: age 60, 55 and 50 for men, and age 55, 50 and 45 for women.
- **Not everyone gets retired exactly at the date of the cutoff age**, as one might retire early, be re-employed without signing a labor contract after retirement, or just simply forget to apply for retirement.

$$\underbrace{depre_i}_{\text{A measure of depression level}} = \alpha_0 + \alpha_1 D_i + \alpha_2 \underbrace{(age_i - c)}_{\text{Age Effect wrt. diff. cutoffs}} + \alpha_3 (age_i - c) Z_i + f(income_i) + \underbrace{u_i}_{\text{Clustered in communities}}$$

A scale of *The Center for Epidemiological Studies-Depression* (CES-D)

- Individuals are asked how often over the past week they experienced symptoms associated with depression, such as restless sleep, poor appetite and feeling lonely.
- The total depression scale ranges from 0 to 30* and a **high score indicates a high level of depression**.
- A cutoff of 8 or 10 is considered depressed.
- We use **several ways of coding** depression for robustness:

  - The sum of scores
  - Log of sum scores
  - A dummy equal to 1 if the score is larger than 8
  - A dummy equal to 1 if the score is larger than 10

**Response options range from 0 to 3 for each item (0 = rarely or none of the time < 1 days; 1 = some or a little of the time 1-2 days; 2 = occasionally or a moderate amount of the time 3-4 days; most or all of the time 5-7 days).*

2. Empirical Design

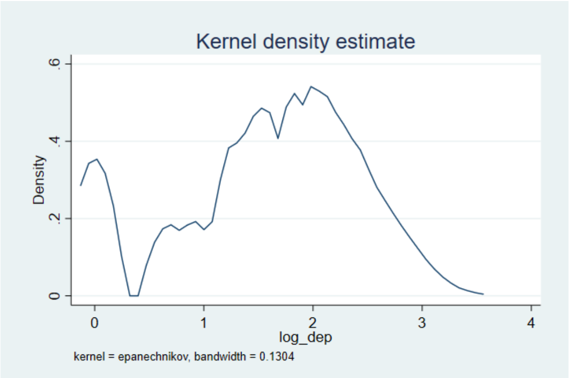
Methodology: Fuzzy RD

- **Three possible cutoffs** in terms of retirement status: age 60, 55 and 50 for men, and age 55, 50 and 45 for women.
- **Not everyone gets retired exactly at the date of the cutoff age**, as one might retire early, be re-employed without signing a labor contract after retirement, or just simply forget to apply for retirement.

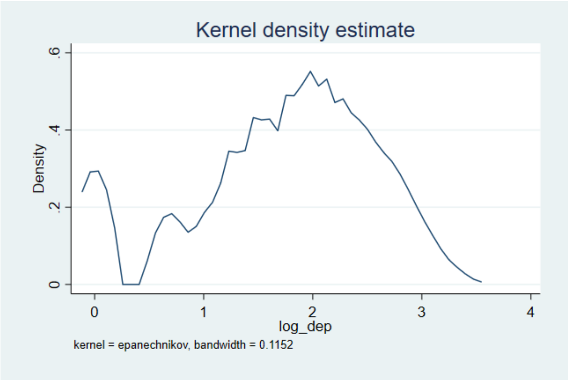
$$\underbrace{depre_i}_{\boxed{\text{A measure of depression level}}} = \alpha_0 + \alpha_1 D_i + \underbrace{\alpha_2 (age_i - c)}_{\text{Age Effect wrt. diff. cutoffs}} + \alpha_3 (age_i - c) Z_i + f(income_i) + \underbrace{u_i}_{\text{Clustered in communities}}$$

A scale of *The Center for Epidemiological Studies-Depression* (CES-D)

- It is hard for Chinese people to talk about depression / admit mental issues.
- People tend to show a positive self-image.
- Abnormal bunches at the 0: interviewees response all “very good” to EACH question.
- We drop out samples with depression score 0.



Male



Female

2. Empirical Design

Methodology: Fuzzy RD

- **Three possible cutoffs** in terms of retirement status: age 60, 55 and 50 for men, and age 55, 50 and 45 for women.
- **Not everyone gets retired exactly at the date of the cutoff age**, as one might retire early, be re-employed without signing a labor contract after retirement, or just simply forget to apply for retirement.

$$\underbrace{depre_i}_{\text{A measure of depression level}} = \alpha_0 + \boxed{\alpha_1} D_i + \alpha_2 \underbrace{(age_i - c)}_{\text{Age Effect wrt. diff. cutoffs}} + \alpha_3 (age_i - c) \boxed{Z_i} + f(income_i) + \underbrace{u_i}_{\text{Clustered in communities}}$$

Possible Drawbacks

- α_1 only specify the effects on the **compliers** it may be only a fraction of the population.
  - Not in China, where the majority are the compliers.

- The IV may **fail to meet the exogeneity requirement** because age has direct effects on physical health.
  - This correlation could be well eliminated if we limit the sample on a short interval around the cutoff.

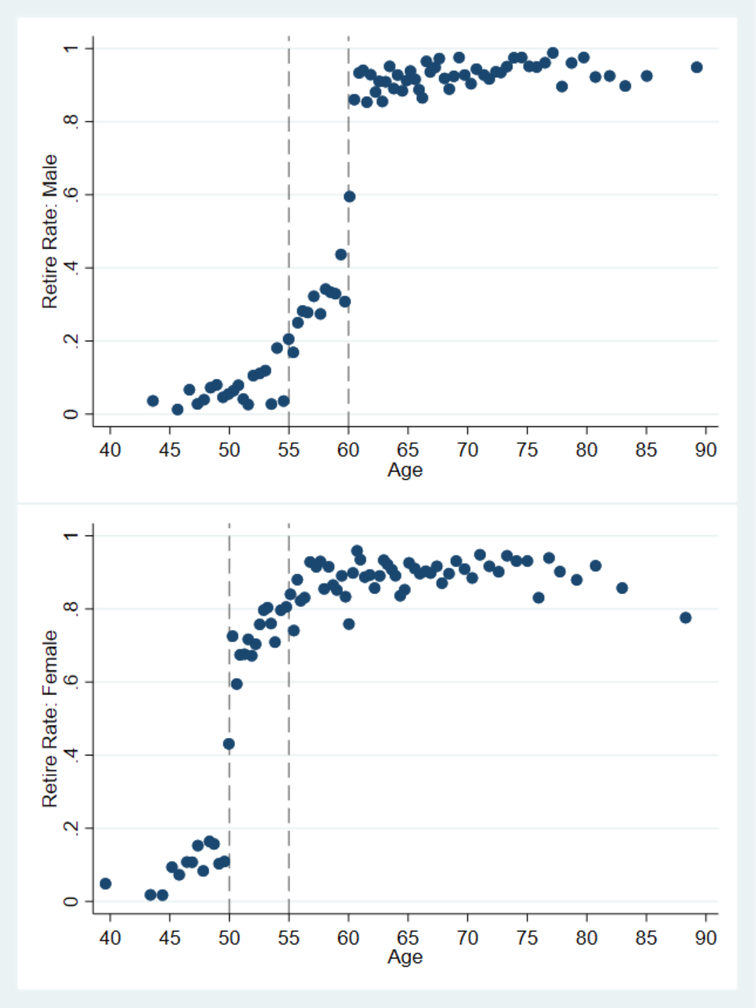
3. Empirical Results

Discontinuity of treatment variable

- First stage regression with difference bandwidths.
- A change **from below cutoff to above cutoff** will **increase the probability of retirement** by 37.5 to 47.7 pct for male.
- 44.4 to 47.4 pct for female.
- In all cases, the square of t-statistics is above 10, which show high relevance between statutory retirement age and retirement status.

VARIABLES	(1) retired	(2) retired	(3) retired	(4) retired
cutoff	0.474***	0.370***	0.474***	0.444***
	(0.0368)	(0.0443)	(0.0535)	(0.0965)
Observations	2,122	2,066	960	339
R-squared	0.375	0.380	0.301	0.257
Bandwidth	5	2	3	1
F statistic	401.6	674.7	151.4	44.52

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1



Male

Female

3. Empirical Results

Continuity of covariables

- Fuzzy RD for males and females subsample resp., using **year of education** and **marriage status** as dependent variable.
- Education for males: although year of education is 1.65 shorter for below cutoff than above cutoff, it is not statistically significant.
- Education for females: the magnitude is small and not statistically significant.
- The probability of marriage also exhibits no significant difference.

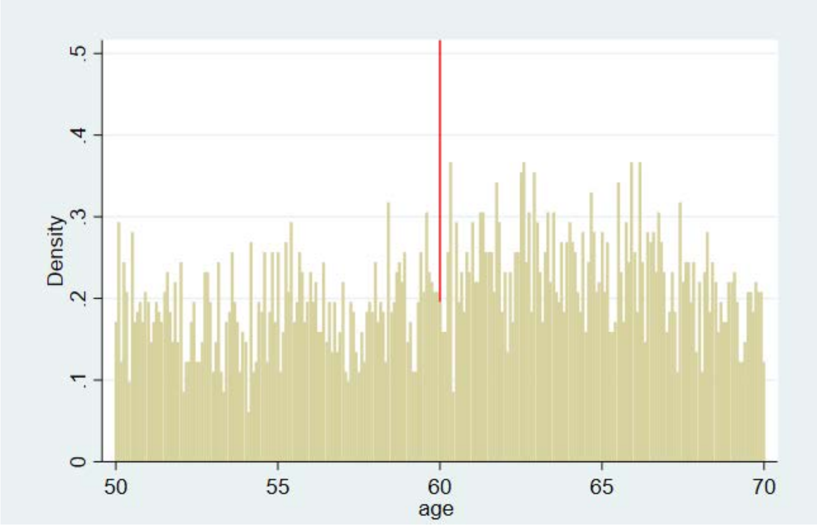
VARIABLES	(1) edu_adj_c	(2) edu_adj_c	(3) marriage	(4) marriage
RD_Estimate	-1.648	0.512	0.122	-0.0488
	(1.487)	(1.082)	(0.0883)	(0.0889)
Observations	3,299	2,121	3,520	3,488
Obser l.t cutoff	367	339	551	526
Obser r.t cutoff	457	462	742	733
Bandwidth	2.457	2.931	3.565	4.725
Order of polynomial	1	1	1	1

Standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

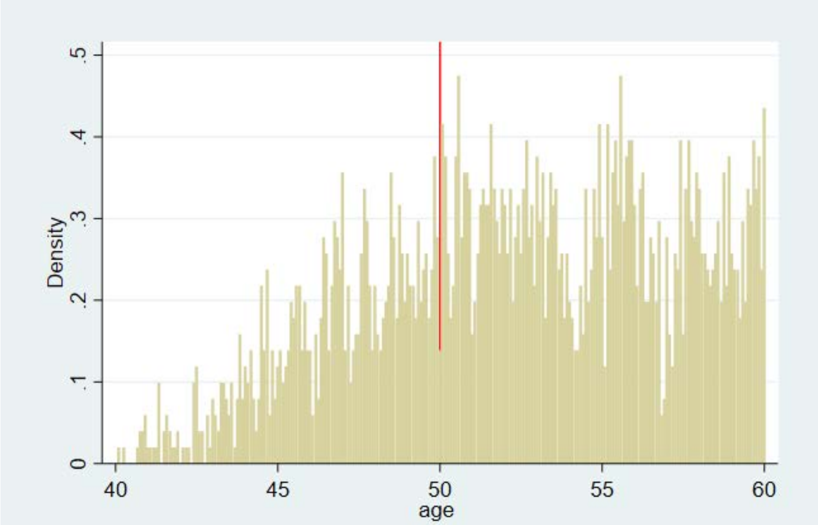
3. Empirical Results

Density of running variable besides cutoff

- No bunching of running variable around cutoff.
- To statistically confirm this, we also use **rddensity** in stata, and get **p-value 0.25 and 0.39** for males and females, respectively.
- No manipulation / misreporting: age is not self-reported by interviewee, instead directly calculated from demographic information from official database.



Male



Female

3. Empirical Results

Baseline results

- First Column: with polynomial of order one.
- Second Column: with polynomial of order two.
- Third Column: add education year & marriage status

VARIABLES	(1) log_dep	(2) log_dep	(3) log_dep
RD_Estimate	0.0310	-0.101	-0.0458
	(0.147)	(0.230)	(0.146)
Observations	3,574	3,574	3,420
Adding controls	No	No	Yes
Obser l.t cutoff	717	771	759
Obser r.t cutoff	943	1014	1000
Bandwidth	5.193	5.661	5.819
Order of polynomial	1	2	1

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Male

VARIABLES	(1) log_dep	(2) log_dep	(3) log_dep
RD_Estimate	0.140	-0.0387	0.0891
	(0.176)	(0.295)	(0.197)
Observations	1,969	1,969	2,924
Adding controls	No	No	Yes
Obser l.t cutoff	421	356	333
Obser r.t cutoff	586	501	455
Bandwidth	4.263	3.449	3.253
Order of polynomial	1	2	1

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Female

Male:
- Small magnitude.
- All not statistically significant.
- Averagely no effect.

Female:
- Small magnitude.
- All not statistically significant.
- Averagely no effect.

3. Empirical Results

Results by education

- 12 years: needed to finish primary school and high school.
- First & Second Column: < 12 years of education.
- Third & Fourth Column: ≥ 12 years of education.

VARIABLES	(1) log_dep	(2) log_dep	(3) log_dep	(4) log_dep
RD_Estimate	0.200	0.443	-0.611	-1.001*
	(0.267)	(0.418)	(0.375)	(0.606)
Observations	1,794	1,794	1,269	1,269
Obser l.t cutoff	209	240	231	354
Obser r.t cutoff	443	515	192	264
Bandwidth	3.539	4.016	3.189	4.707
Order of polynomial	1	2	1	2

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Male

VARIABLES	(1) log_dep	(2) log_dep	(3) log_dep	(4) log_dep
RD_Estimate	0.288	0.296	-0.0124	-0.0829
	(0.306)	(0.377)	(0.244)	(0.318)
Observations	998	998	971	971
Obser l.t cutoff	179	249	189	217
Obser r.t cutoff	233	334	282	321
Bandwidth	3.423	5.117	3.624	4.560
Order of polynomial	1	2	1	2

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Female

Male:
- Positive with < 12 years edu
- Negative with ≥ 12 years edu
- Statistically significant with ≥ 12 years edu

Female:
- Positive with < 12 years edu
- Negative with ≥ 12 years edu
- Not statistically significant

3. Empirical Results

Results by marriage status

- First & Second Column: unmarried / separated / divorced / widowed.
- Third & Fourth Column: married with spouse present.

VARIABLES	(1) log_dep	(2) log_dep	(3) log_dep	(4) log_dep
RD_Estimate	0.0663	0.110	-0.101	-0.242
	(0.316)	(0.382)	(0.252)	(0.437)
Observations	270	270	2,793	2,793
Obser l.t cutoff	43	48	386	458
Obser r.t cutoff	50	55	535	652
Bandwidth	3.307	3.667	3.200	3.866
Order of polynomial	1	2	1	2

Standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Male

- Positive for the unmarried
- Negative for the married
- Not statistically significant

VARIABLES	(1) log_dep	(2) log_dep	(3) log_dep	(4) log_dep
RD_Estimate	-0.382	-0.540	0.142	-0.0494
	(0.617)	(0.835)	(0.210)	(0.312)
Observations	275	275	1,694	1,694
Obser l.t cutoff	41	58	312	307
Obser r.t cutoff	69	89	425	411
Bandwidth	3.028	4.352	3.457	3.353
Order of polynomial	1	2	1	2

Standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Female

- Negative for the unmarried
- Ambiguous for the married
- Not statistically significant

THANKS

Q&A

Kai Chiu Yang & Tinghuan Liu

Program Evaluation and Causal Inference