

# CHINA HEALTH AND RETIREMENT LONGITUDINAL STUDY (CHARLS)

## PROGRESS AND RECENT FINDINGS

YAFENG WANG

PEKING UNIVERSITY

5<sup>th</sup> International SHARE User Conference

13, Nov, 2015

# OUTLINE

- ▶ Progress
  - ▶ Motivation
  - ▶ Introduction to CHARLS
  - ▶ Sampling
  - ▶ CHARLS Data and Questionnaire
- ▶ Findings
  - ▶ Health and Support
  - ▶ Vignettes and Self Reported Health
  - ▶ Employment Trends

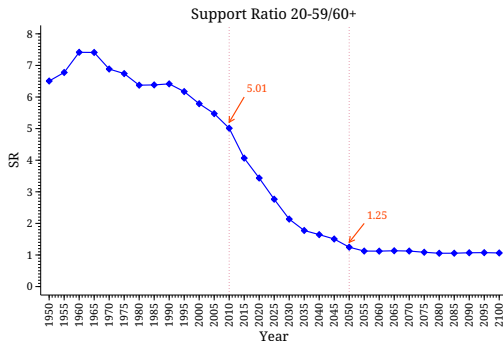
# THE SPEED OF POPULATION AGING

Number of years for population age 65+ to increase from 7% to 14%

Developed Countries		Developing Countries	
France (1865-1980)	115	Azerbaijan (2000-2041)	41
Sweden (1890-1975)	85	Chile (1998-2015)	27
Australia (1938-2011)	73	China (2000-2026)	26
US (1944-2013)	69	Jamaica (2008-2033)	25
Canada (1944-2009)	65	Tunisia (2008-2032)	24
Hungary (1941-1994)	53	Sri Lanka (2004-2027)	23
Poland (1966-2013)	47	Thailand (2003-2025)	22
UK (1930-1975)	45	Brazil (2011-2032)	21
Spain (1947-1992)	45	Colombia (2017-2037)	20
Japan (1970-1996)	26	Singapore (2000-2019)	19

- China has one of the fastest ageing population in the world.

# SUPPORT RATIO



- ▶ One of the major challenges of population aging is the shrinkage of working population relative to those in retirement
- ▶ The current retirement age of 60 will quickly become unsustainable



# CHARLS

- ▶ CHARLS: HRS-type biennial panel of nationally representative of population over age 45
- ▶ <http://charls.ccer.edu.cn/en>

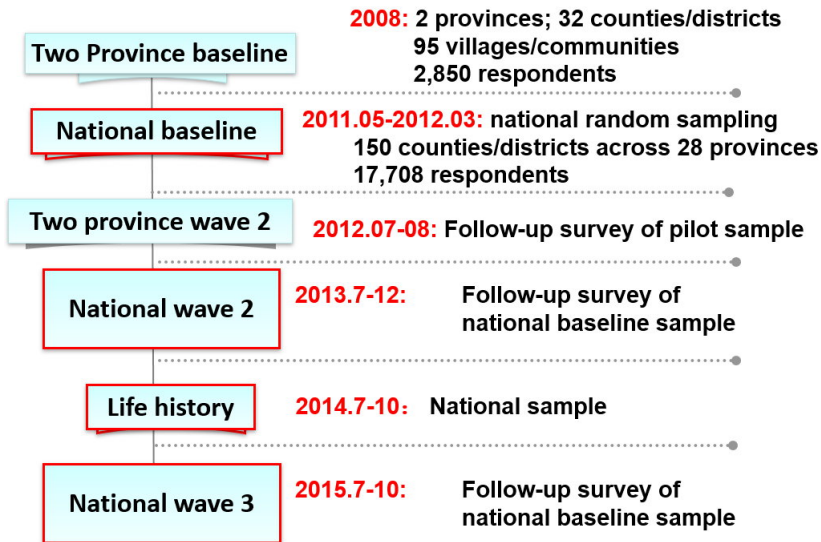
# HRS FAMILY

- ▶ International family of survey data
  - ▶ US: HRS, 1992
  - ▶ UK: ELSA, 2002
  - ▶ Europe: SHARE, 2004
- ▶ Asia
  - ▶ Japan: JSTAR, 2007
  - ▶ Korea: KLoSA, 2008
  - ▶ Indonesia: IFLS 1998
  - ▶ India: LASI, 2014
- ▶ Others: Mexico, Brazil, Ireland, Australia
- ▶ CHARLS is modeled after the international series and is internationally comparable

# FUNDING AGENCIES

- ▶ NSFC
  - ▶ Regular project (2008-2011)
  - ▶ Key international collaboration project (2009-2013)
  - ▶ Data infrastructure project (2011-)
- ▶ Peking University
- ▶ NIA/NIH
- ▶ World Bank

# TIMELINE



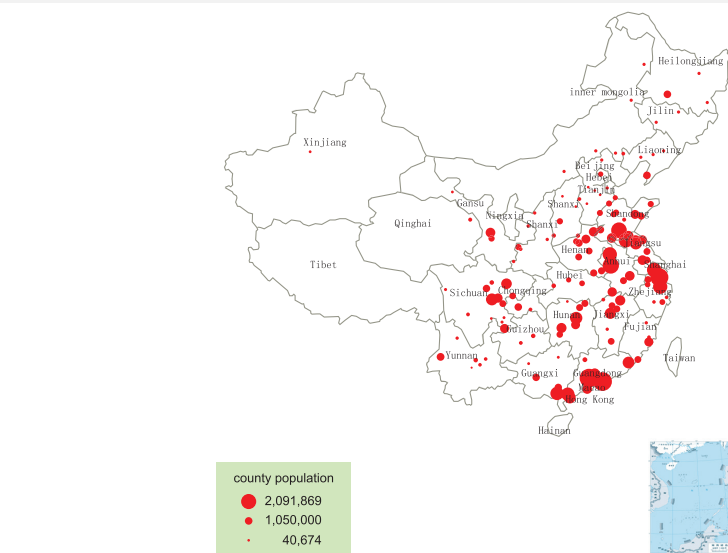
# USERS

**TABLE:** Distribution of Users by Country

Country	Freq.	Percent	Cum.
China	10,338	87.15	87.15
USA	837	7.06	94.21
UK	177	1.49	95.7
Australia	63	0.53	96.23
Canada	50	0.42	96.65
Singapore	59	0.5	97.15
France	29	0.24	97.4
Japan	57	0.48	97.88
Netherlands	37	0.31	98.19
Germany	36	0.3	98.49
Korea	34	0.29	98.78
Sweden	18	0.15	98.93
Others	127	1.07	100
Total	11,862	100	

# NATIONAL BASELINE SAMPLING

- ▶ Multi-stage PPS random sampling
  - ▶ Counties (stratification: by regions, urban/rural, pc GDP): 150
  - ▶ Villages/communities: 450
  - ▶ Household sampling frame created out of the mapping/listing operation
  - ▶ One household member 45 and older randomly chosen, plus the spouse
- ▶ Sample size
  - ▶ Households: 10,257; Persons: 17,708



# MAPPING & LISTING

- ▶ There is no pre-existing sampling frame at the village level
- ▶ No list of mailing addresses
- ▶ Lists of residents exist in some communities, not all
- ▶ Some lists are outdated



# CHARLS-GIS

- ▶ Computerize the mapping/listing operation
  - ▶ Use google earth map as base
  - ▶ Edit for changes
  - ▶ Complete listing of dwellings within a building
  - ▶ Records GPS and photo of each building
  - ▶ Records photo of each door of sample dwelling



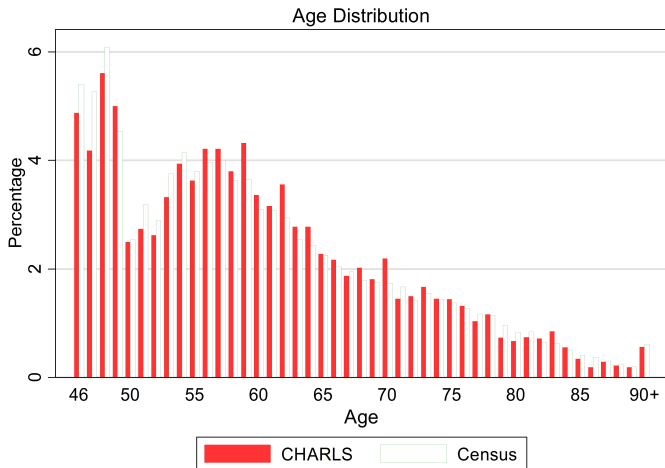


# RESPONSE RATES

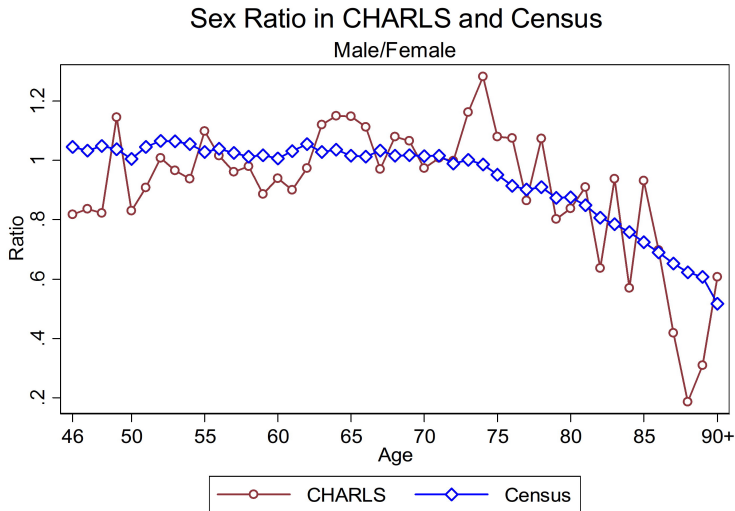
TABLE: Response Rates

		2008	2011	2012	2013		2014	
					Cross Section	Panel	Cross Section	Panel
Response rate (%)	Total	84.82	80.51	94.02	82.63	88.30	78.27	85.76
	Rural	89.69	94.15	97.18	91.74	92.18	91.96	91.44
	Urban	79.33	68.63	90.04	72.20	82.61	66.69	77.43
No. of households	Total	1,570	10,257	1,494	10,629	9,022	12,250	8,826
	Rural	879	6,033	861	6,340	5,547	7,243	5,540
	Urban	691	4,224	633	4,289	3,475	5,007	3,286
No. of respondents	Total	2,683	17,708	2,381	18,264	15,196	20,547	14,481
	Rural	1,501	10,537	1,370	10,950	9,439	12,297	9,187
	Urban	1,182	7,171	1,011	7,314	5,757	8,250	5,294

# CHARLS VS. CENSUS



# CHARLS vs. CENSUS



# QUESTIONNAIRE

## Household and Individual:

- ▶ Cover screen
- ▶ Household roster
- ▶ Demographics
- ▶ Family structure/transfer
- ▶ Health status, functioning and care giving
- ▶ Biomarkers
- ▶ Blood Samples
- ▶ Health care and insurance
- ▶ Work, retirement and pension
- ▶ Income and consumption
- ▶ Assets
- ▶ Housing conditions
- ▶ Exit (followup)
- ▶ Verbal Autopsy (followup)

## County and Community:

- ▶ Community survey
- ▶ Policy survey

# BIOMARKERS

- ▶ Anthropometric measurements:
  - ▶ height, weight, waist circumference, lower right leg length and upper arm length
- ▶ Lung capacity, grip strength, timed sit and stand, timed walk, balancing
- ▶ Hypertension

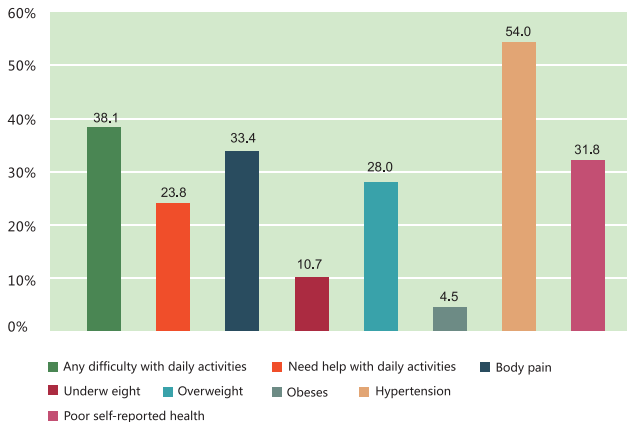


# BLOOD SAMPLES

- ▶ National Baseline is collected under a sister project from NSFC by the China CDC
- ▶ National Baseline blood sample collected from ~ 11,804 individuals, while 2015 followup is about 13,357
  - ▶ Response Rate: ~ 66.7%
- ▶ Blood sample as part of a sister survey of CHARLS to be analyzed for:
  - ▶ Complete Blood Count (hemoglobin, etc.)
  - ▶ R01 supp: CRP, HbA1c, Lipids (total, LDL and HDL cholesterol, Triglycerides), Cystatin C

# HEALTH INDICATORS

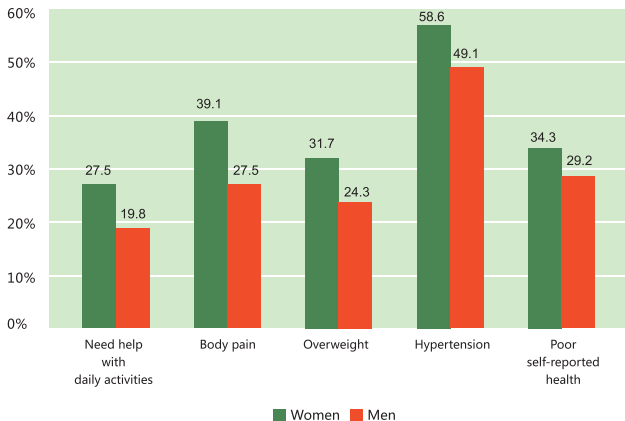
FIGURE: Physical Health Status of the Elderly



► Elderly health is worrisome

# GENDER GAP

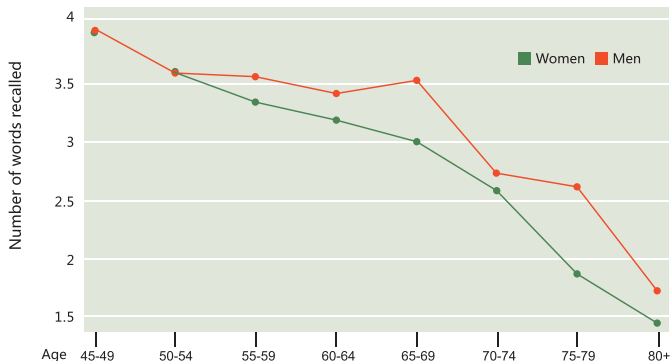
FIGURE: Physical Health Status of the Elderly: by Gender



- ▶ Gender gap seems larger in China
- ▶ Gender inequality goes beyond socio-economic areas

# COGNITION

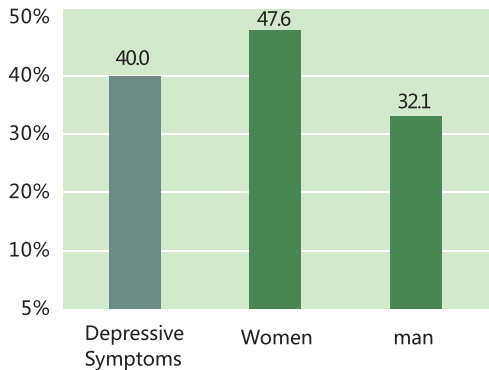
FIGURE: Cognition Ability by Gender and Age Group



- ▶ Cognition declines sharply with age; Female worse
- ▶ This makes the elderly susceptible to fraudulence targeting the elderly
- ▶ Caring for those with the Alzheimers disease will be a big challenge

# MENTAL HEALTH

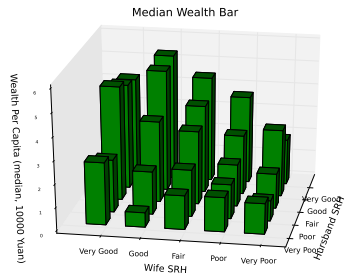
FIGURE: Depressive Symptoms



- ▶ A large number of the elderly display higher levels of depressive symptoms; women worse

# HEALTH AND SES

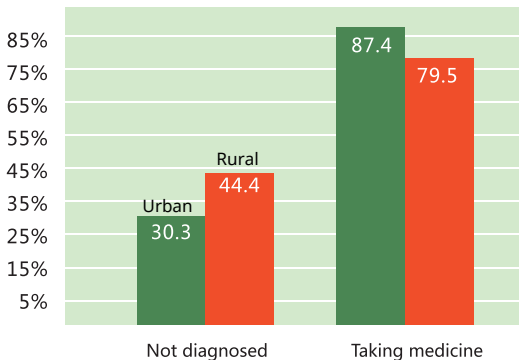
FIGURE: Wealth Per Capita and Self-Reported Health



- ▶ Strong SES gradients of wealth
- ▶ Causality may go both directions

## MEDICAL SERVICE

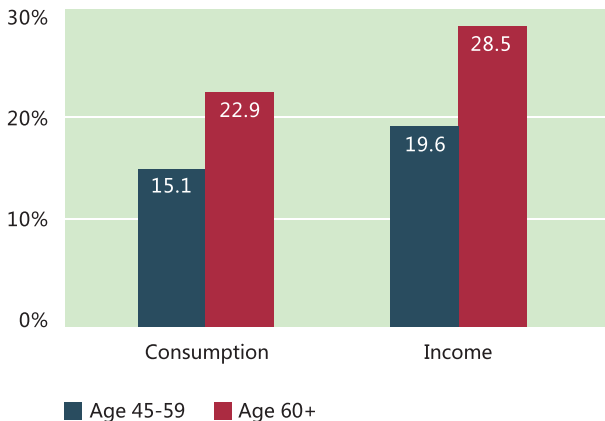
**FIGURE:** Underdiagnosis of Hypertension and Share of Taking Medication if Diagnosed



- ▶ Large urban-rural gap in awareness of hypertension
- ▶ Treatment is lower in rural as well

# POVERTY

FIGURE: Consumption and Income Poverty Rates by Age Group

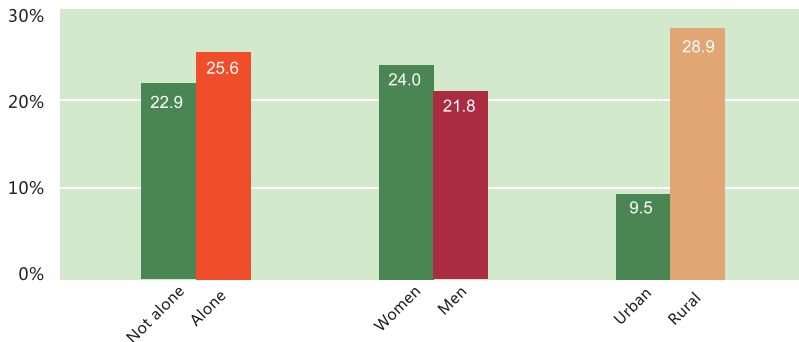


- ▶ Expenditure is better measure of poverty
- ▶ The elderly is more vulnerable



# POVERTY

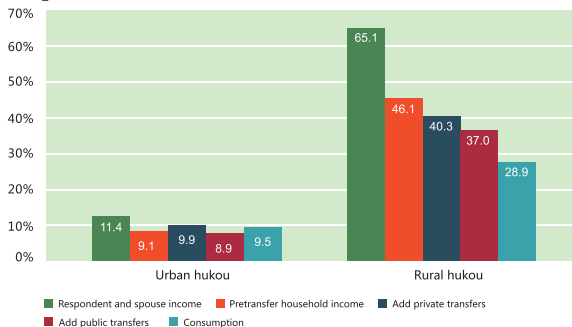
**FIGURE:** Elderly Consumption Poverty Rates by Whether Living Alone, Gender, and Hukou



- ▶ Rural elderly are the poorest
- ▶ Living alone, older women are poorer

# THE ROLE OF FAMILY AND GOVERNMENT

**FIGURE:** Elderly Poverty Rates Using Different Definitions of Per Capita Income and Expenditure



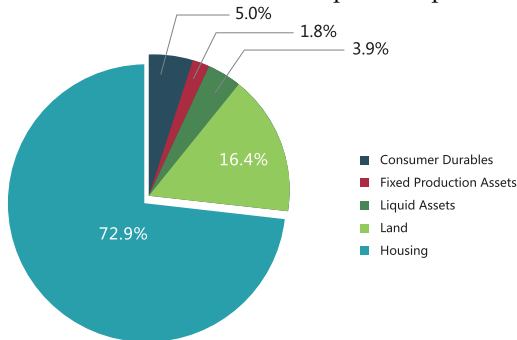
- ▶ For rural elderly, living with children is most effective in reducing poverty; transfers from children, government and other sources are important too
- ▶ For urban elderly, co-residence has a small role in reducing poverty. Children are net burdens; government has a small role

# WEALTH

- ▶ In rural areas, median per capita wealth is 18,000 yuan; can support 5 years of median per capita expenditure
- ▶ In urban areas, median per capita wealth is 71,000 yuan; can support 8 years of median per capita expenditure

# COMPOSITION OF WEALTH

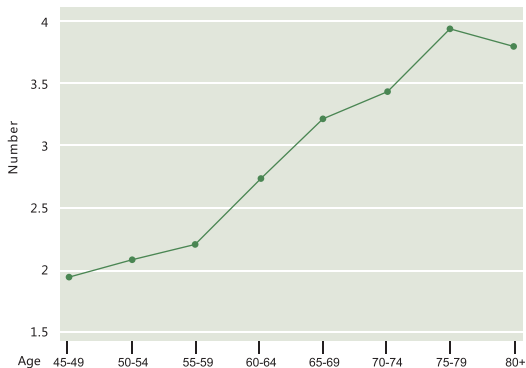
FIGURE: Median Wealth Per Capita Composition



- ▶ Housing makes up 73% of wealth (urban 86.5%, rural 68.5%). Financial innovations is needed to utilize housing as a source of finance for old age support
- ▶ Land is 22% of rural wealth. The value of land can increase if land can be sold

# NUMBER OF CHILDREN

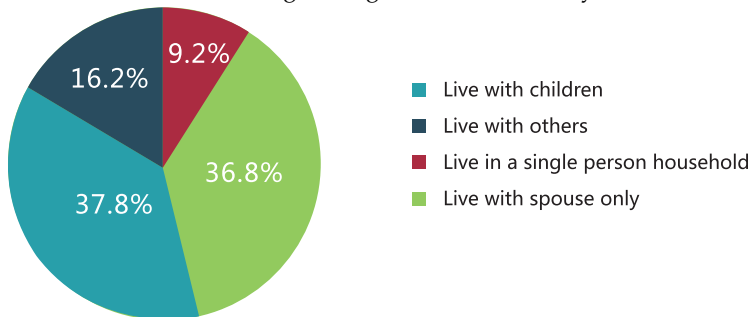
FIGURE: Number of Living Children by Age Group



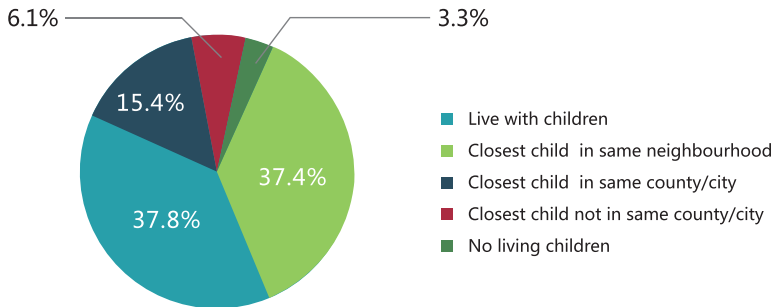
- ▶ The number of children that the elderly can rely on will reduce sharply
- ▶ In ten years the number of children for 65-69 cohort will decline by  $\frac{1}{3}$

# LIVING ARRANGEMENT

FIGURE: Living Arrangement of the Elderly



► 46% of elderly will be “empty-nest”



- ▶ Living alone does not necessarily mean having no help. Most empty-nest elderly have children nearby. Very few (6.1%) elderly's children live out of the county/city
- ▶ In the future, should encourage the migration of the elderly parents

# FAMILY TRANSFERS

**FIGURE:** Transfer between Elderly and Non-coresident Children

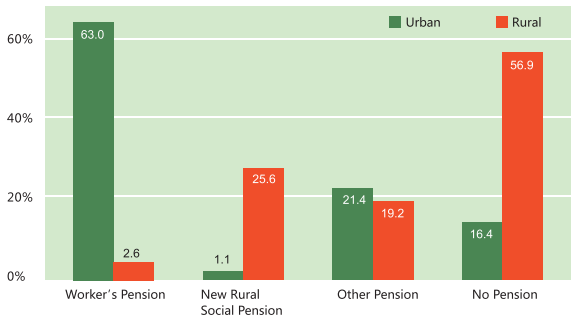
Percent receiving transfers from non-coresident children	46.9
Percent giving transfers to non-coresident children	19.0
Percent receiving transfers from children if not living with children and have children	53.3
Median positive net transfers received by those who receive positive net transfers from non-coresident children (yuan)	1,700
Median net transfers received as share of elderly household expenditures (for those receiving positive net transfers) (%)	37.3

- ▶ Contrary to developed countries, the direction of transfers is upward
- ▶ Among parents who receive support from children, 37% of expenditure is financed by children



# PENSION COVERAGE

FIGURE: Pension Coverage of the Elderly by Hukou



- ▶ Urban hukou 83.6% rural 43.1%
- ▶ New rural social pension has benefitted 25.6% rural elderly

# PENSION BENEFITS

FIGURE: Pension Benefits of the Elderly

Pension Scheme	Individuals Receiving the Pension (%)	Median Annual Pension Benefit (yuan)	Median Share of Expenditure (%)
Pension Subsidy to the Oldest Old	3.6	720	19.1
New Rural Social Pension Insurance	18.2	720	21.0
Urban and Other Residents' Pension	14.4	1,200	38.7
Firm Basic Pension	10.7	18,000	192.9
Firm Supplemental Pension	1.2	13,200	169.7
Government or Institutions' Pension	10.9	24,000	242.2
Commercial Pension	0.3	14,400	106.1
Other Pension	2.7	9,600	150.7

- ▶ Firm pension is indeed lower than that of the government, but both types of pension pay for two persons' living expenses on average
- ▶ The least generous is rural pension, which covers only 21% of one persons expenses

# SUMMARY

- ▶ Chinese older population face many health challenges
- ▶ Elderly poverty is serious. Although social security is expanding, the current level of support is low
- ▶ Future support from family will be more difficult; need to study how the function of family can be maintained. Need to study the migration of the elderly parents.
- ▶ Need to study how to extend the working lives of urban workers so as to reduce the burden to the family and the society

# VIGNETTES



James P. Smith, Yafeng Wang and Yaohui Zhao. Vignettes and Self-Reported Health among the Elderly in China

- ▶ Using vignettes information to adjust the state-dependent reporting bias in Self Reported Health
- ▶ Find significant heterogeneity in both location and scale

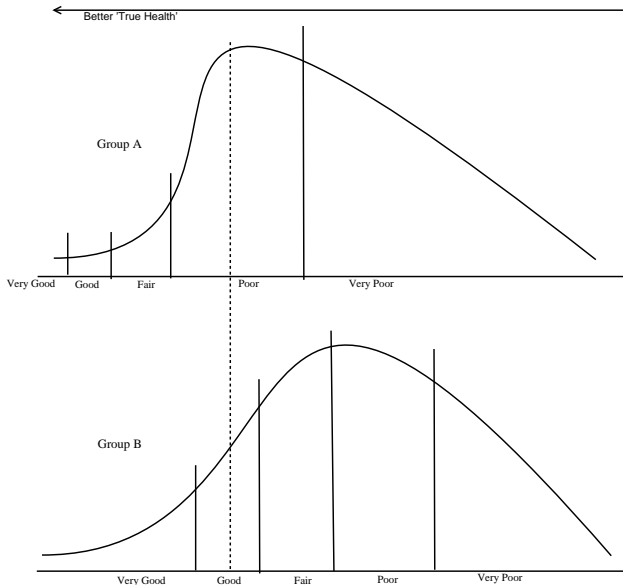
# SRH IN CHARLS

**QUESTION DESIGNED TO DETERMINE THE SRH IN CHARLS** I have some questions about your health. Would you say your health is very good, good, fair, poor or very poor?

1. Very Good
2. Good
3. Fair
4. Poor
5. Very Poor

# MOTIVATION

- ▶ Adequate estimates of the role of SES on health do critically depend on the appropriateness of the health measures used.
  - ▶ Objective vs. Subjective Health Measures (self-reported health)
- ▶ It is sometimes argued that the mapping of 'true health' into self-reported health (SRH) categories may vary with respondent characteristics:
  - ▶ State-dependent reporting bias (Kerkhofs and Lindeboom, 1995)
  - ▶ Response category cut-point shift (Murray et al. 2001)
  - ▶ Scale of reference bias (Groot 2000)
  - ▶ Differential item functioning (DIF) (King et al. 2004)



## RESEARCH QUESTION

- ▶ If our concern is about that how  $X$  affect the self-reported health, then the usual ordered choice model (e.g., OPROBIT) is appropriate.
- ▶ But if the concern is about that how  $X$  affect the actual health level, we should take account of the reporting heterogeneity.
  - ▶ Since the thresholds in the ordered choice model is a function of  $X$ , the usual ordered choice model estimation is not consistent.
  - ▶ The vignettes approach proposed by Gary King, Murray, Salomon, and Tandon (2004, American Political Science Review) can solve such comparability problem.
- ▶ Kapteyn, Smith and Soest (2007, AER): the self-reported work disability in the Netherlands is higher than the United States, when using vignettes to adjust the threshold heterogeneity, the resulting work disability in the two countries are the same (statistically).



# VIGNETTERS IN CHARLS

## A TYPICAL SURVEY QUESTION OF VIGNETTE IN CHARLS: A/B

falls asleep easily at night, but two nights a week he/she wakes up in the middle of the night and cannot go back to sleep for the rest of the night. In the last month, how much difficulty did A/B have with sleeping, such as falling asleep, waking up frequently during the night, or waking up too early in the morning.

1. None
2. Mild
3. Moderate
4. Severe
5. Extreme

# INTUITION

- ▶ A vignette question describes the health of a hypothetical person and then asks the respondent to evaluate that person's health **on the same scale** used for a self-report on their own health.
- ▶ So
  - ▶ Given the same actual health of the hypothetical person in the vignette question, differences in respondents' evaluations must due to thresholds heterogeneity.
  - ▶ Given the estimation of thresholds heterogeneity based on vignettes, we can consistently estimate the role of X on actual health.

## MODEL: SELF-ASSESSMENT COMPONENT

- ▶ The actual health level for respondent  $i$  is  $Y_i^*$ , a continuous unidimensional variable (corresponding to the order of the response categories, a higher value means worse health status).
- ▶ The actual health level  $Y_i^*$  is assumed to be a linear function of observed covariates  $X_i$  and an independent normal random effect  $e_i$ :

$$Y_i^* = X_i\beta + e_i \quad (1)$$

$$e_i \rightsquigarrow N(0, \omega^2) \quad (2)$$

- ▶ The reported survey response category is  $y_i$  and is generated by:

$$y_i = k \text{ if } \tau_i^{k-1} \leq Y_i^* < \tau_i^k \quad (3)$$

- ▶  $\tau_i$  vary over the respondents, and can be modeled as a function of covariates  $X_i$  and unknown parameter vector  $\gamma$ .
- ▶ Two specifications for this function with  $\tau_i^0 = -\infty$ ,  $\tau_i^K = \infty$ .

- ▶ Linear case:

$$\tau_i^1 = X_i\gamma^1; \tau_i^k = \tau_i^{k-1} + X_i\gamma^k \quad (k = 2, \dots, K-1) \quad (4)$$

- ▶ Nonlinear case:

$$\tau_i^1 = X_i\gamma^1; \tau_i^k = \tau_i^{k-1} + \exp(X_i\gamma^k) \quad (k = 2, \dots, K-1) \quad (5)$$

- ▶ **Identification Failure.**

## VIGNETTE COMPONENT

- ▶ The actual health level for the hypothetical person in vignette  $j$  is  $\theta_j$  ( $j = 1, \dots, J$ ), measured on the same scale as  $Y_i^*$  and the  $\tau$ 's.
- ▶ Respondent perceives  $\theta_j$  with normal error: ([Vignette Equivalence](#))

$$Z_{ij}^* \rightsquigarrow N(\theta_j, \sigma_j^2) \quad (6)$$

which represents respondent  $i$ 's unobserved assessment of the health level of vignette  $j$ .

- ▶ The respondent turns the continuous  $Z_{ij}^*$  into a categorical answer to the survey question  $z_{ij}$  via the same mechanism in self-assessment component: ([Response Consistency](#))

$$z_{ij} = k \text{ if } \tau_i^{k-1} \leq Z_{ij}^* < \tau_i^k \quad (7)$$

- ▶ The vignettes are used to identify the parameters in the thresholds, and given these thresholds values, using self-reports one can identify  $\beta$ .
- ▶ Combining the two components results the [Hierarchical Ordered Probit Model \(HOPIT\)](#)

## MORE HETEROGENEITY

- ▶ Heterogeneity in the actual health level:

$$e_i \rightsquigarrow N(0, \omega_i^2) \quad (8)$$

$$\omega_i = \exp [X_i' \delta_s] \quad (9)$$

- ▶ Heterogeneity in the perceiving the hypothetical person's health:

$$Z_{ij}^* \rightsquigarrow N(\theta_j, \sigma_{ij}^2) \quad (10)$$

$$\sigma_{ij} = \exp [X_i' \delta_v + \delta_j] \quad (11)$$

- ▶ We call the resulting model as **Heterogeneous HOPIT model (HHOPIT)**.
- ▶ Using LR test to select the 'true' model among HOPIT and HHOPIT.

# ESTIMATION

- ▶ The two components share several parameters and so should be estimated together.
- ▶ Maximum Likelihood Estimation.
- ▶ Location and Scale Normalization: No intercept in the actual health level equation,  $\omega_i$  and  $\sigma_{ij}$
- ▶ There is no proof of the global concavity of the HHOPIT likelihood function
  - ▶ We use both Newton and Genetic algorithms.
  - ▶ Two-step estimation as starting value.

TABLE: Summary Statistics

Variable	Vignettes Respondent	Other Respondent	All
Self Reported Health	3.00	3.01	3.01
Female	0.54	0.53	0.53
Rural	0.77	0.78	0.78
Age	58.76 (9.82)	59.14 (10.01)	59.05 (9.97)
Illiterate	0.26	0.27	0.27
Literate	0.18	0.18	0.18
Primary	0.23	0.22	0.22
Middle School	0.20	0.21	0.20
High School and Above	0.13	0.12	0.13
Married	0.86	0.88	0.87
Ln PCE	8.73 (0.87)	8.73 (0.87)	8.73 (0.87)
Observations	3604	12588	16192

Standard deviations in parentheses



TABLE: Regression Result 1

VARIABLES	OPROBIT		HOPIT							
	SRH		SRH		CutI		Self Scale		Vig Scale	
	coef	se	coef	se	coef	se	coef	se	coef	se
Constant					-2.372***	0.157				
Rural	0.044*	0.027	0.012	0.035	0.044	0.032				
Female	0.074***	0.020	0.085***	0.027	-0.051**	0.025				
Age55-64	0.068***	0.023	0.038	0.030	0.006	0.028				
Age65-74	0.133***	0.029	0.073*	0.039	-0.083**	0.036				
Age75+	0.152***	0.042	0.242***	0.060	0.079	0.055				
Literate	-0.034	0.030	0.068	0.048	-0.019	0.058	0.000	0.041	-0.013	0.039
Primary	-0.071**	0.029	0.030	0.043	0.131***	0.048	-0.130***	0.037	-0.189***	0.034
Middle School	-0.179***	0.032	-0.142***	0.043	0.158***	0.050	-0.215***	0.038	-0.242***	0.035
High School and Above	-0.218***	0.039	-0.212***	0.051	0.175***	0.058	-0.290***	0.044	-0.280***	0.042
Married	-0.025	0.031	-0.013	0.040	-0.026	0.037				
Ln PCE	-0.156***	0.013	-0.106***	0.016	0.023	0.015				
Has Hypertension and Know	0.282***	0.029	0.281***	0.040	-0.038	0.037				
Has Hypertension but Don't Know	-0.066**	0.027	-0.056	0.036	-0.032	0.033				
Controlled Hypertension	0.284***	0.032	0.330***	0.043	-0.017	0.039				
Lung Disease	0.381***	0.030	0.272***	0.041	0.003	0.037				
Liver Disease	0.345***	0.048	0.146**	0.066	-0.193***	0.065				
Heart Disease	0.497***	0.030	0.372***	0.043	-0.108***	0.040				
Digestive Disease	0.378***	0.023	0.227***	0.032	-0.155***	0.031				
Arthritis	0.363***	0.021	0.275***	0.029	-0.082***	0.026				
Log Likelihood	-15799.461				-39534.80715					
Observations	13,344				13,344					

- ▶ Significant heterogeneity in both location and scale
- ▶ Male, High Educated, Richer, Healthier will be more optimistic in evaluating self-reported health

TABLE: Regression Result 2

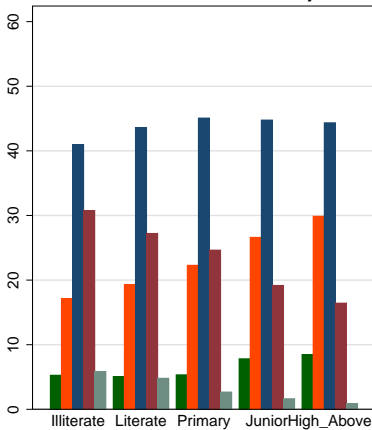
VARIABLES	OPROBIT		HOPIT							
	SRH		SRH		Cut1		Self Scale		Vig Scale	
	coef	se	coef	se	coef	se	coef	se	coef	se
Constant					-2.697***	0.139				
Rural	-0.002	0.025	-0.039	0.029	0	0.026				
Female	0.139***	0.019	0.141***	0.023	-0.043**	0.02				
Age55-64	0.130***	0.022	0.1***	0.026	0.009	0.023				
Age65-74	0.247***	0.028	0.195***	0.034	-0.027	0.03				
Age75+	0.234***	0.039	0.331***	0.052	0.076*	0.046				
Literate	0.013	0.029	0.066	0.045	0.023	0.056	-0.021	0.047	-0.034	0.044
Primary	-0.057**	0.028	-0.001	0.039	0.142***	0.045	-0.177***	0.04	-0.24***	0.038
Middle School	-0.161***	0.03	-0.13***	0.038	0.202***	0.045	-0.278***	0.04	-0.267***	0.038
High School	-0.194***	0.036	-0.168***	0.043	0.26***	0.049	-0.396***	0.044	-0.335***	0.041
Married	-0.02	0.03	-0.028	0.036	-0.013	0.032				
Ln PCE	-0.171***	0.012	-0.126***	0.014	0.023*	0.012				
Famine Intensity	0.001	0.002	0.007**	0.003	0.009***	0.002				
Log Likelihood	-18255.629		-43272.42263							
Observations	14,298		14,298							

- ▶ Those experienced famine will be more optimistic in evaluating self-reported health

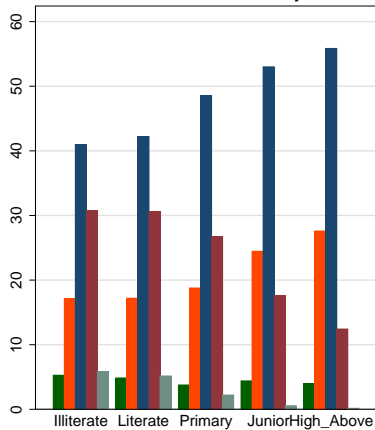
# SIMULATION

- ▶ Using simulation to show the effect of threshold heterogeneity.
- ▶ Comparing two simulated self-reported health based on the HHOPIT estimates:
  - ▶ All the individuals use same  $X$  in threshold functions except for a subset of  $X$ , say  $X_s$ .
  - ▶ All the individuals use same  $X$  in threshold functions
- ▶ The comparison will show the effect of threshold heterogeneity in  $X_s$

Distribution of SRH without Adjustment



Distribution of SRH with Adjustment



# SUMMARY

- ▶ We find significant thresholds heterogeneity in self-reported health among the elderly in China
  - ▶ Male, High Educated
  - ▶ Those Richer, Healthier, Experienced Famine
  - ▶ will be more optimistic in evaluating self-reported health
- ▶ There is heterogeneity in perceiving the hypothetical person's health
  - ▶ Higher educated, less variation
- ▶ Empirical work based on SRH should take account of comparability problem induced by thresholds heterogeneity.

# EMPLOYMENT



Yafeng Wang, Yaohui Zhao. Employment Trends in China

# QUESTIONS

## Policy Question:

- ▶ Is there a potential for raising employment in the future?

## Research Question:

- ▶ What has been the recent trend of employment/retirement?
  - ▶ The past trend is unknown and not studied
- ▶ What explains the trend?
- ▶ But, no nationally representative survey is available to researchers for such a study

# CHARLS LIFE HISTORY SURVEY 2014

- ▶ Retrospective survey using event calendar method
- ▶ Employment history each job lasted 6 months or more
- ▶ Reconstruct employment rate for particular age groups in past years
- ▶ Hukou status is “then”



# VALIDATE LIFE HISTORY DATA

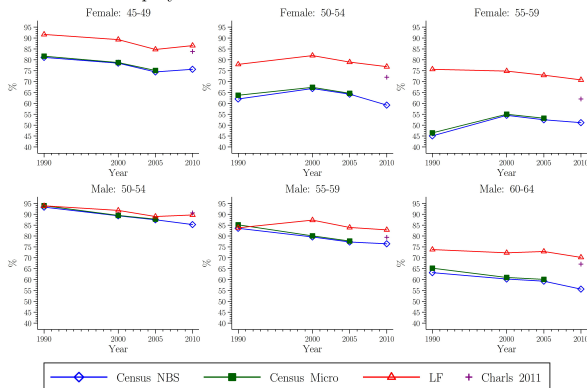
- ▶ Compare with official census estimates
  - ▶ 1990, 2000, 2005 mini census
- ▶ Compare with estimates from census micro data
  - ▶ 1990, 2000, 2005 mini census
- ▶ Separate comparisons for urban and rural estimates with micro-data based census data only

# DEFINITIONAL DIFFERENCES IN EMPLOYMENT

- ▶ Census:
  - ▶ Did you work for pay at least one hour last week?
- ▶ CHARLS:
  - ▶ Did you work at least 10 days in agriculture in the past year?
  - ▶ If not, did you work for pay at least one hour last week?
- ▶ CHARLS is expected to produce higher numbers due to seasonality of agriculture

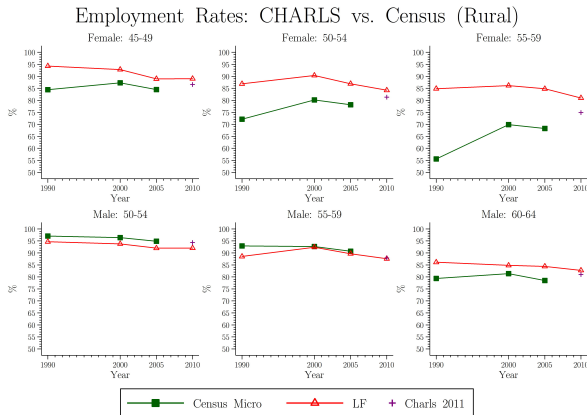
# CHARLS vs. CENSUS

Employment Rates: CHARLS vs. Census



- ▶ Despite the gap, CHARLS trends closely mimic that from census
- ▶ Because urban response rates are lower and LFP are lower too, CHARLS estimates over-estimate participation. With proper weights this gap may reduce

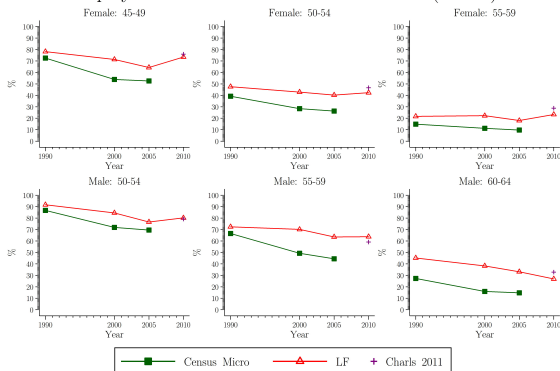
# CHARLS vs. CENSUS: RURAL



- ▶ CHARLS has higher employment rates than census for women, very close in number for men

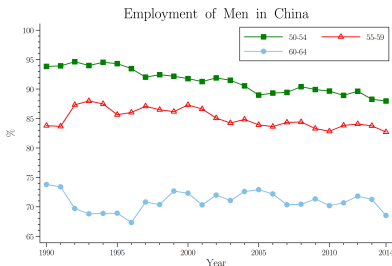
# CHARLS vs. CENSUS: URBAN

Employment Rates: CHARLS vs. Census (Urban)



- ▶ Census seems to under-report urban employment. Unless probed, urban people tend to be shy in reporting self-employment activities
- ▶ Overall, CHARLS life history captures the most recent past 15 years quite well

# EMPLOYMENT TREND



Source: CHARLS Life History

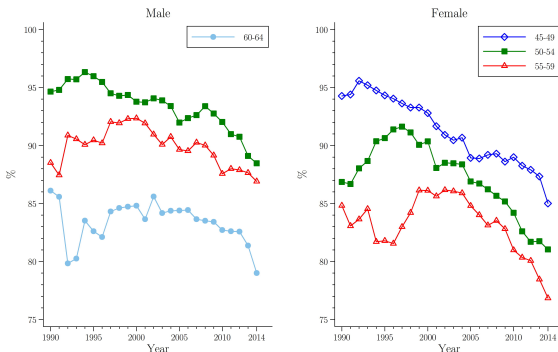


Source: CHARLS Life History

- ▶ There is an overall declining trend in employment rate for both men and women, but interesting patterns exist within various age cohorts.
- ▶ Employment of men aged 60-64 declined in first half of 1990s then rebounded back.
- ▶ Employment of women aged 55-59 declined in first half of 1990, rebounded in second half of 1990, declined thereafter
- ▶ That of women aged 50-54 increased in the 1990s then declined

# RURAL TREND

Rural Employment in China

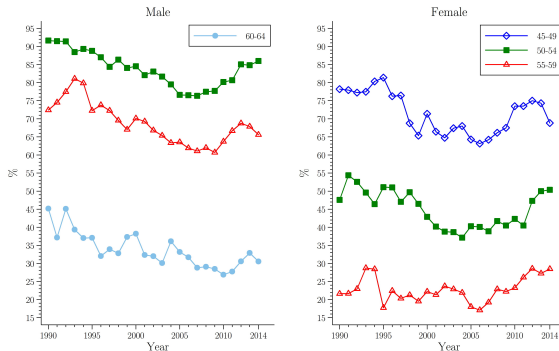


Source: CHARLS Life History

- ▶ There are important and interesting trends in rural employment.
- ▶ Employment rose in the 1990s and declined quite sharply in the 2000s

# URBAN TREND

Urban Employment in China



Source: CHARLS Life History

- ▶ Urban trends are opposite that of rural ones, declined first then stopped
- ▶ The reversing trend started with women around 2005 and men around 2010



Thank you !



王松文Kevin  
renren.com/236040081

