



Mortality Trends and Variation in China

Yong Cai caiyong@unc.edu

University of North Carolina, Chapel Hill

May 15, 2019

Harnack Haus, Berlin

Satellite meeting

Mortality data and methodological approaches in
estimating mortality in developing countries

Sponsored by Max Planck Institute for Demographic Research



Overview

- **Background**
- **Data sources & challenges**
- **Mortality trends and variation**



Socioeconomic background

- **Economic growth**
 - **Changes in nutrition and consumption**
- **Social transformation**
 - **Individual awareness and life style**
- **Establishment of “new” social welfare system**
 - **Healthcare and delivery**
 - **Public health investment**



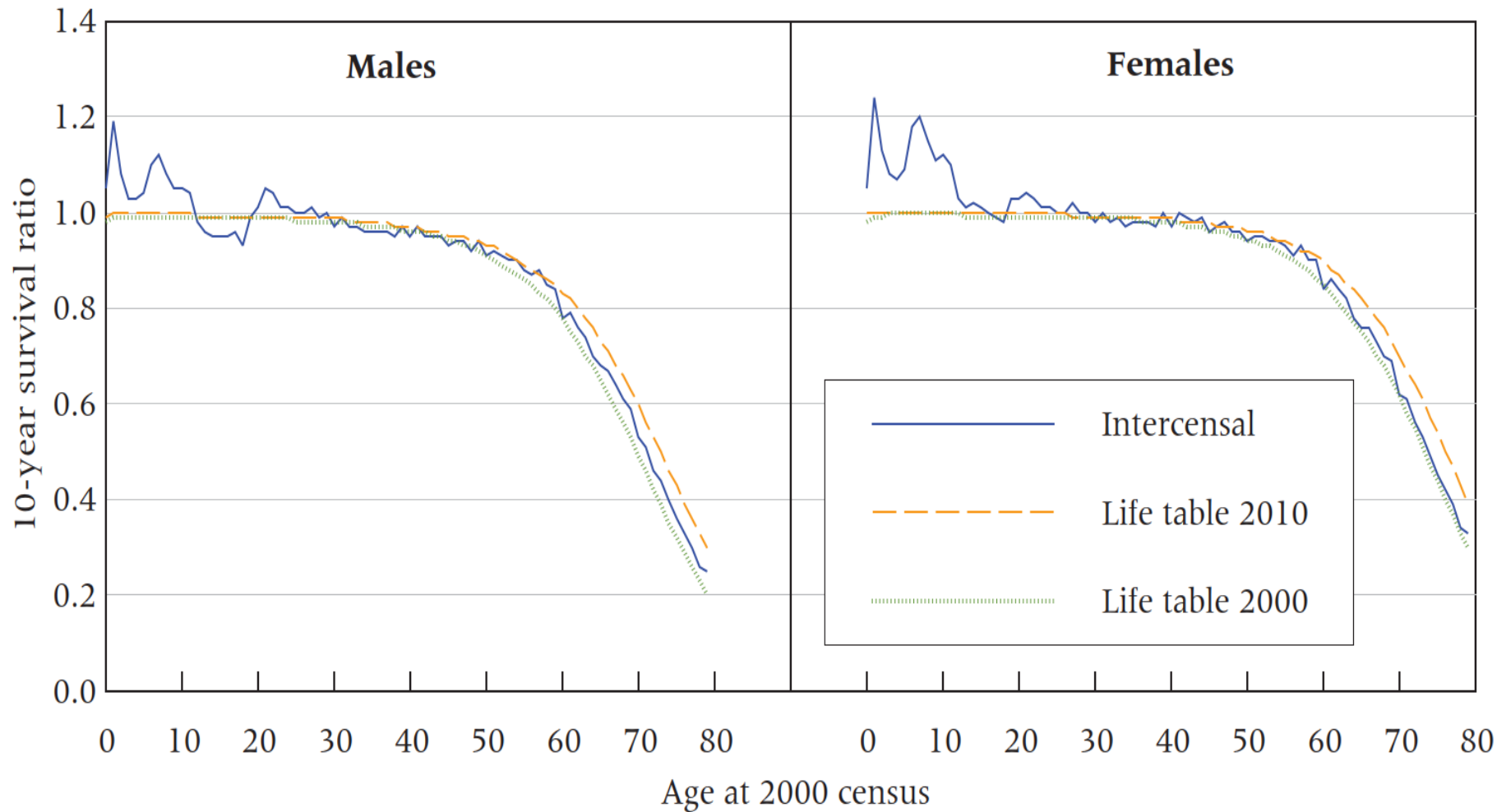
Data sources and challenges

- **Main sources**
 - Population census and surveys
 - Health surveillance systems
 - Death registration
 - Household registration (Hukou)
 - Social pension system
- **Main challenges**
 - **Data quality**
 - Underreporting
 - Migration
 - Government evaluation
 - **Data availability**
 - Segmentation
 - Undocumented adjustment



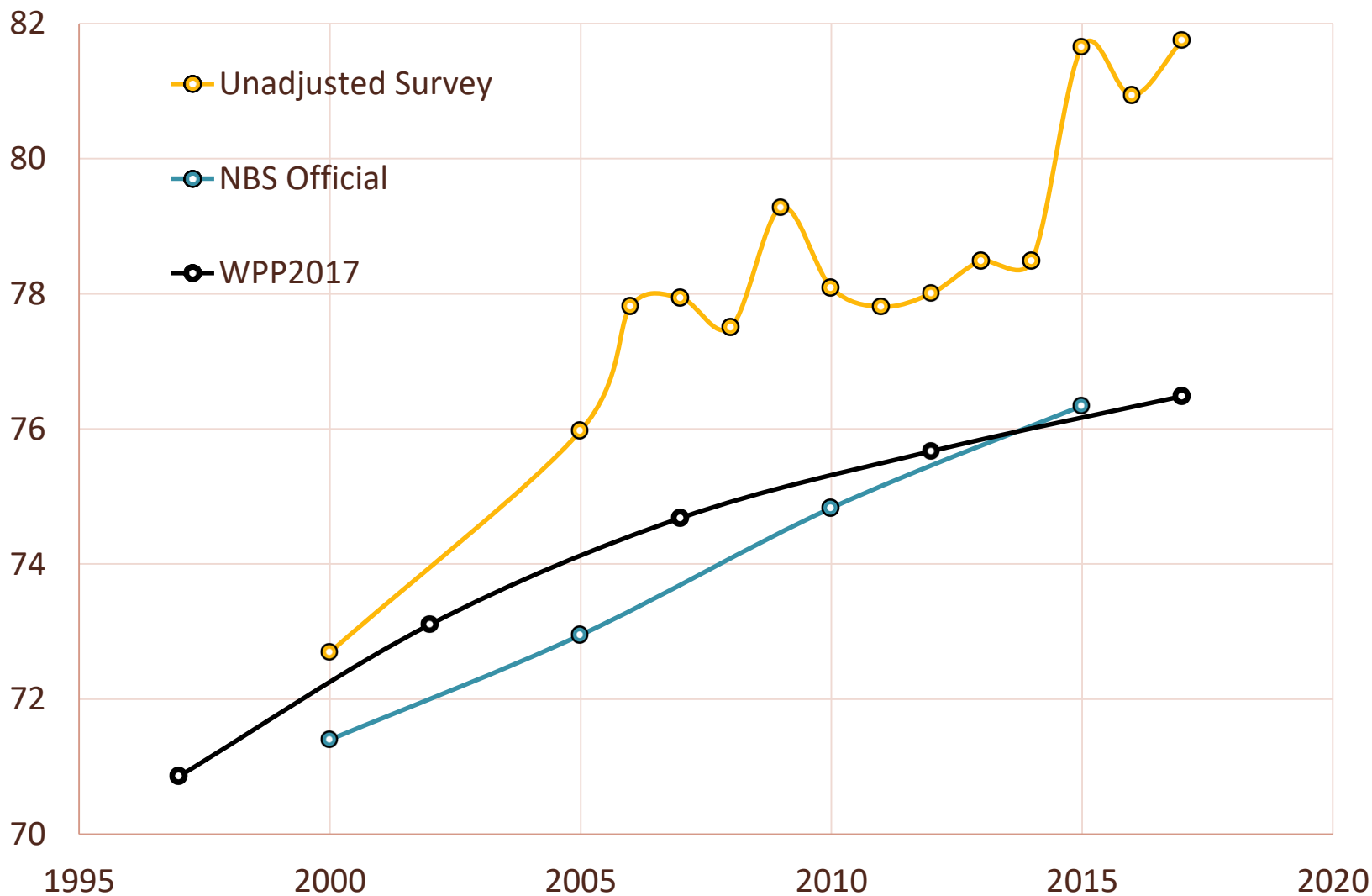
Chinese data: good and bad

FIGURE 2 Intercensal survival ratio by sex and birth cohort, China 2000–2010



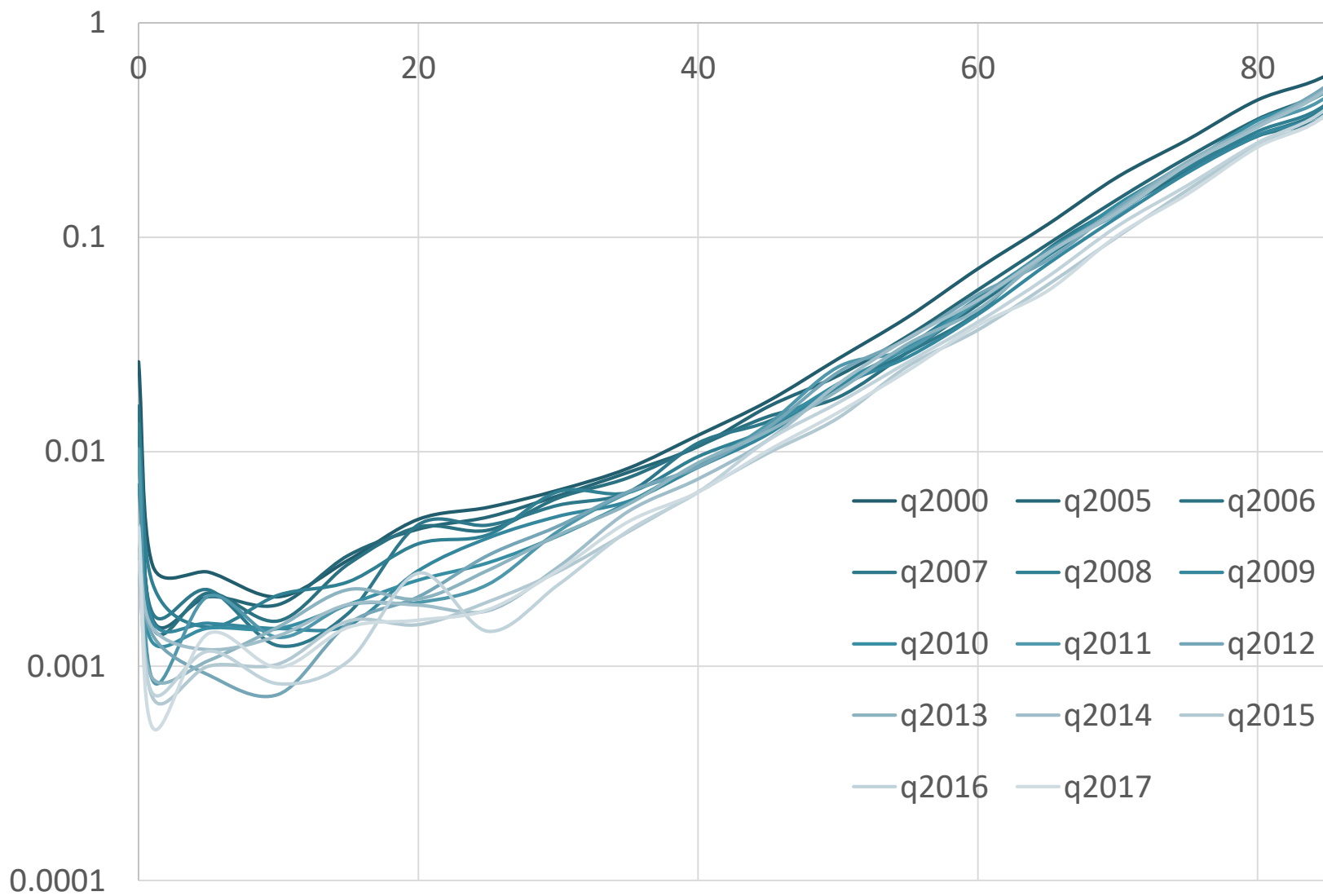


Mortality trends: e_0^t



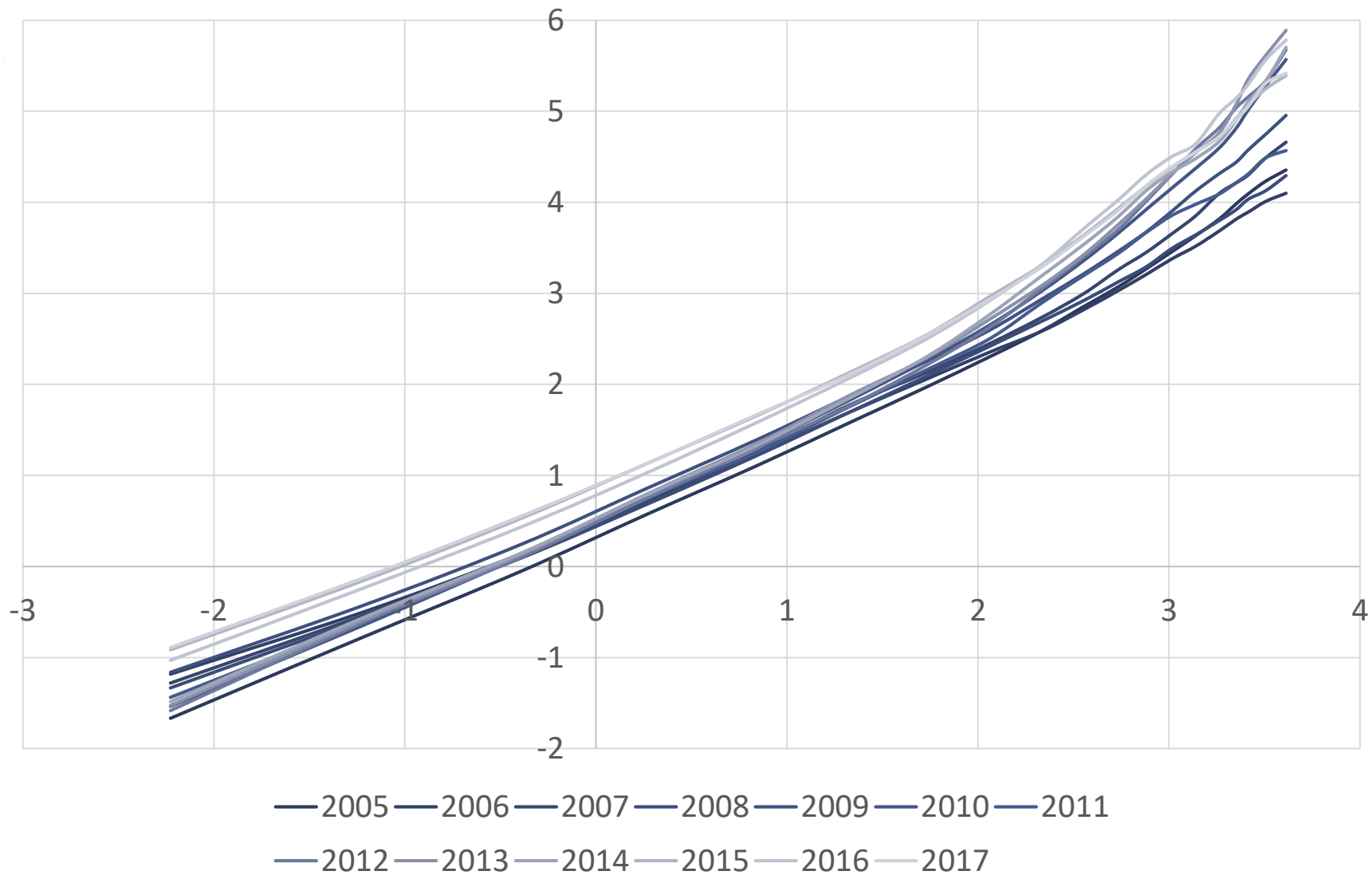


Mortality trends: $q_{\underline{x}}^t$



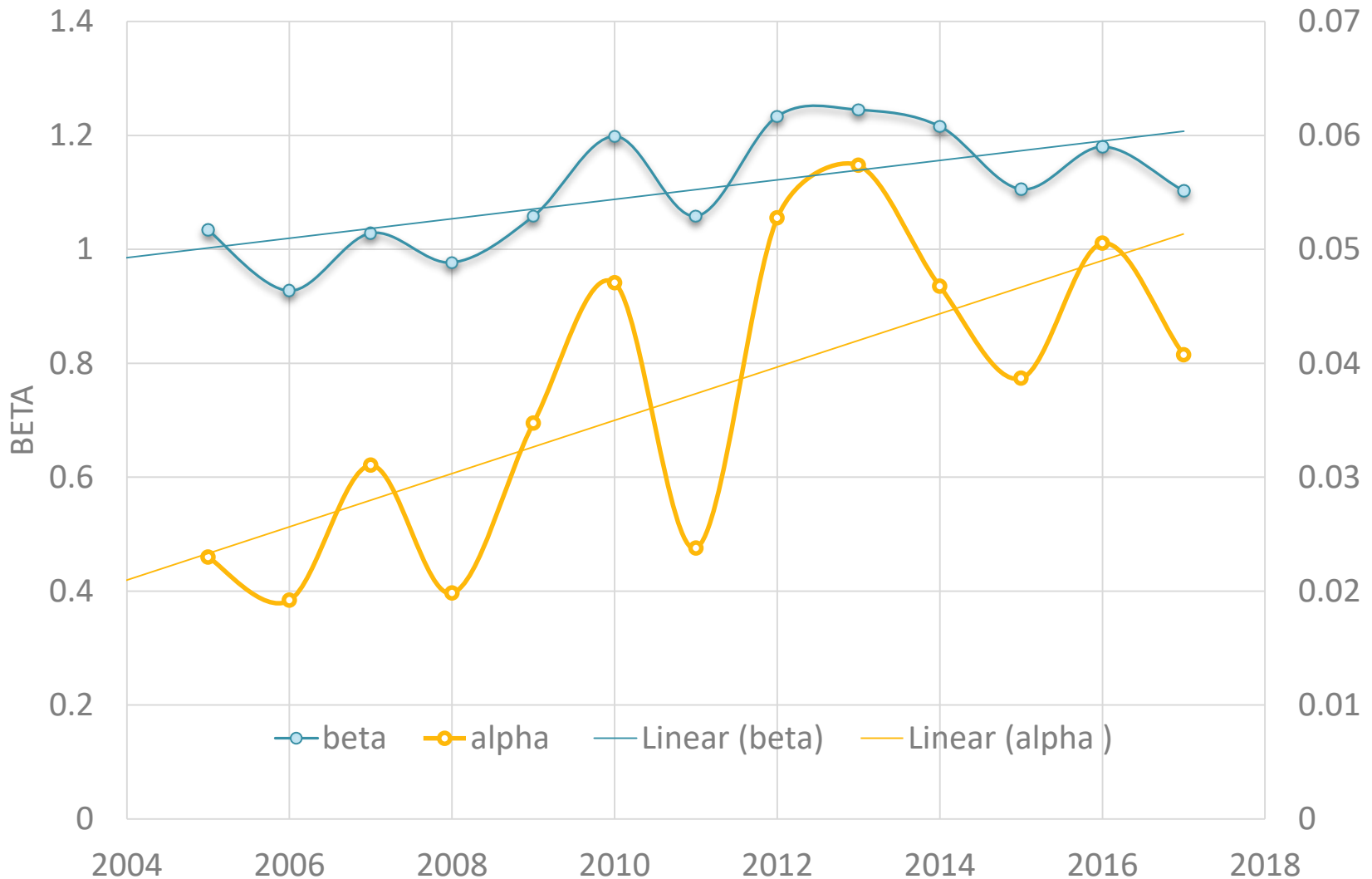


Mortality trends: $\lg t(l_x^t)$ compared to 2000



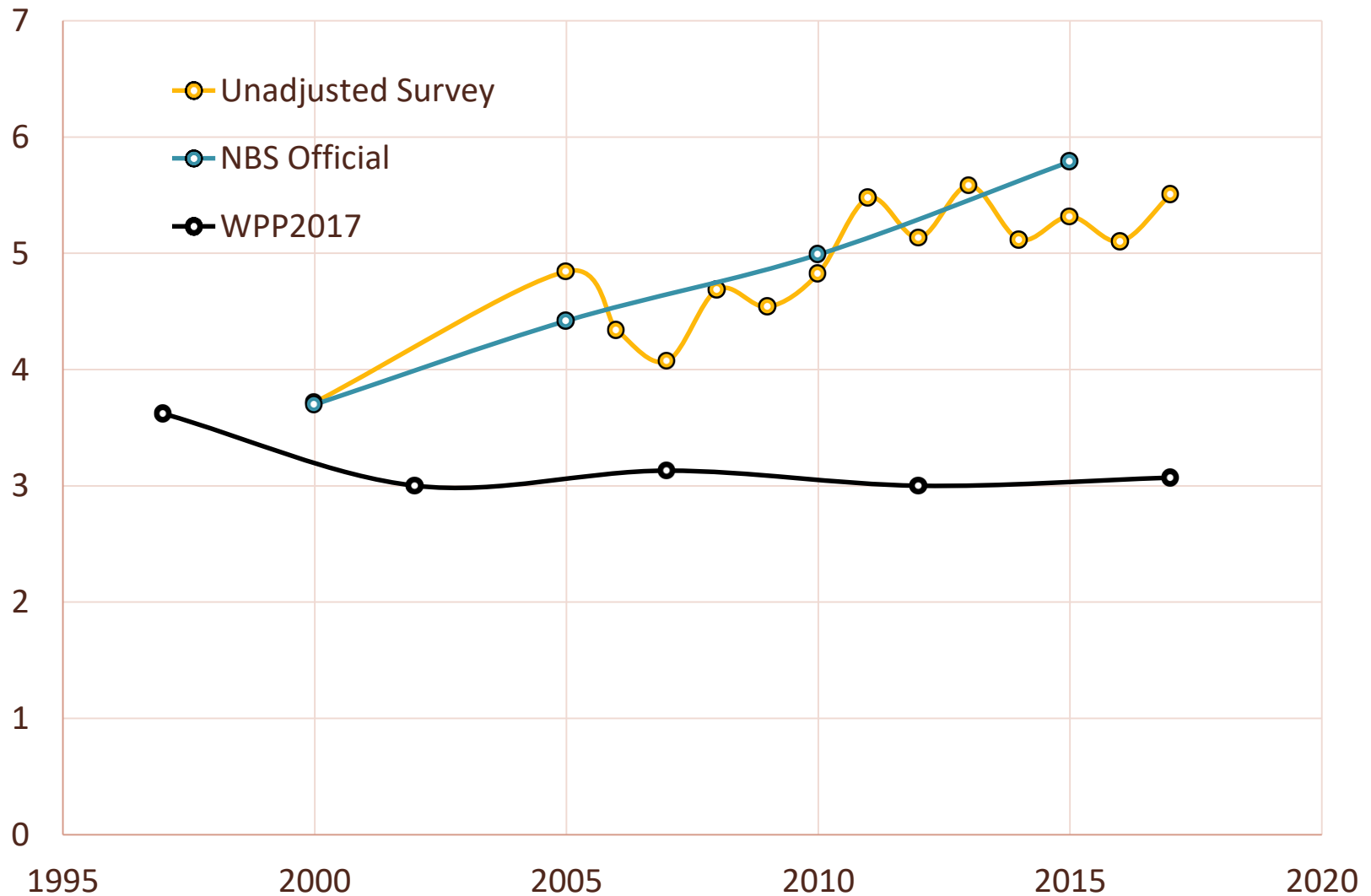


Mortality trends: $\lg t(I_x^t)$ compared to 2000



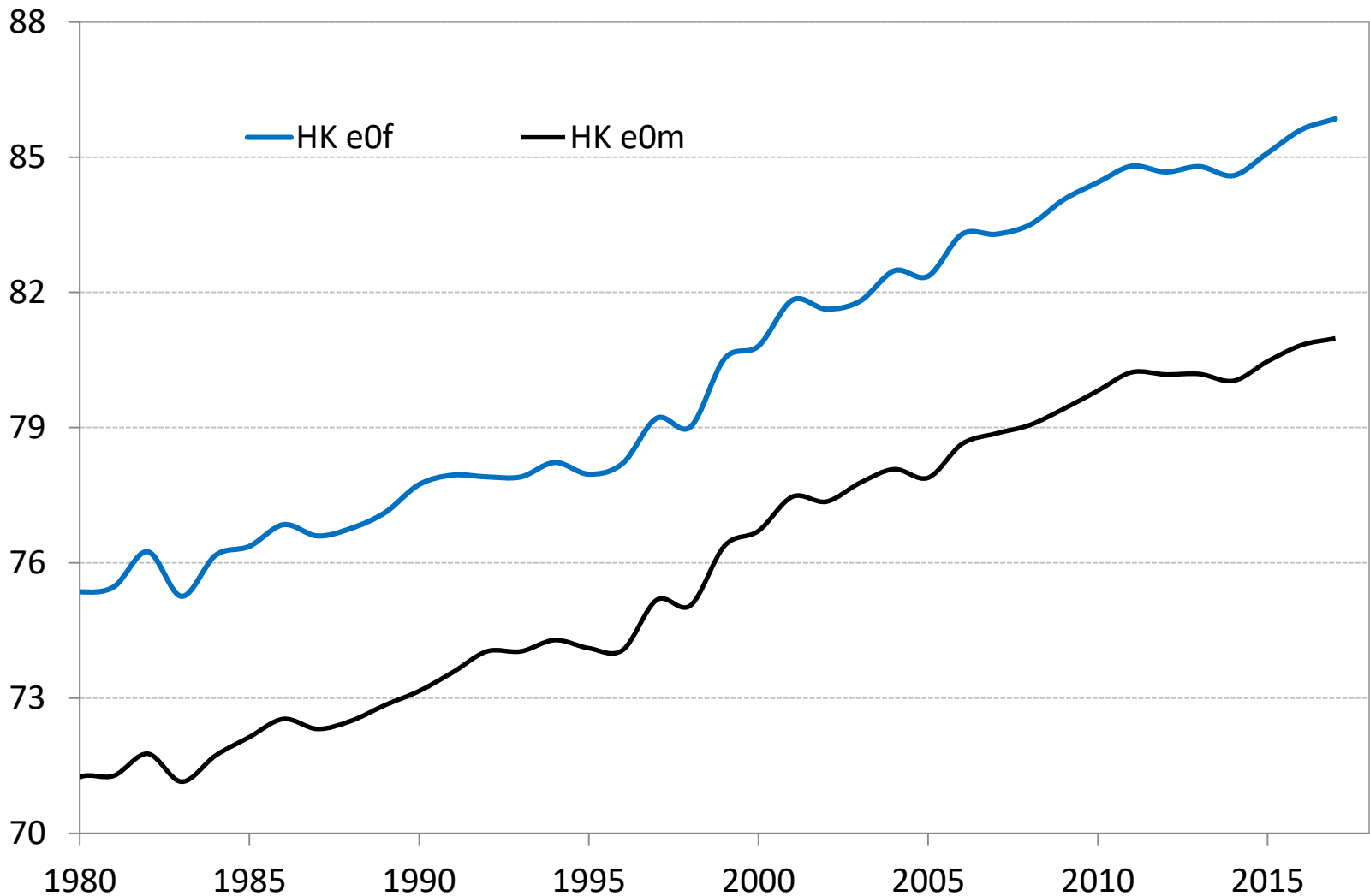


Mortality trends: gender gap $e_0^f - e_0^m$



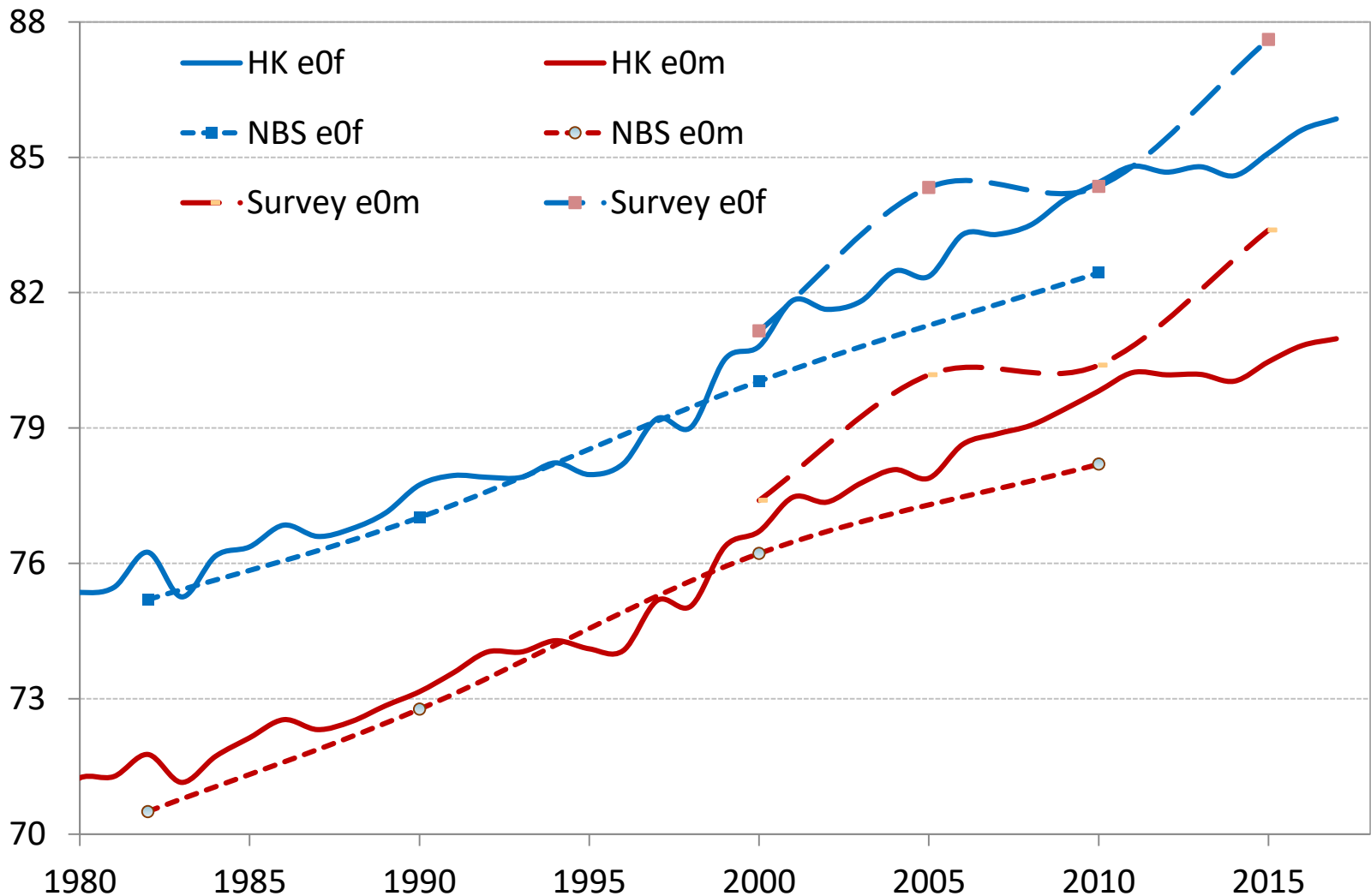


Morality trends: Shanghai e_0



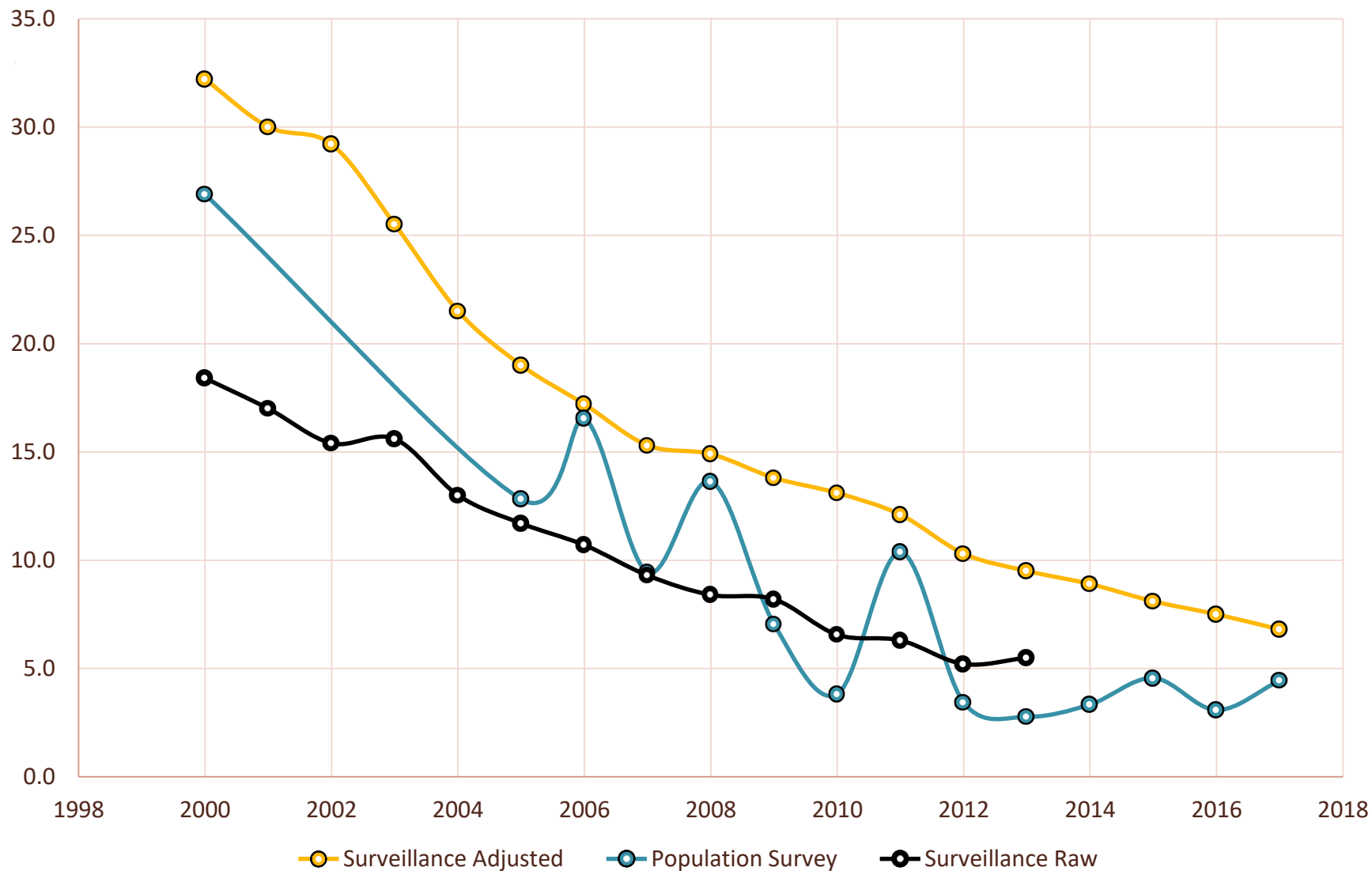


Morality trends: Shanghai e_0





Mortality trends: infant mortality rate





Mortality trends: basic observations

- **Rising life expectancy fits well with the general pattern of mortality decline as it was driven by multiple factors happening in China: socioeconomic development, investment in public health, change of life style...**
- **Mortality underreporting is a constant problem, and will remain to be.**
- **While overall mortality decline can be collaborated with other data sources, ascertain of mortality level requires careful cross-examination with other population data.**

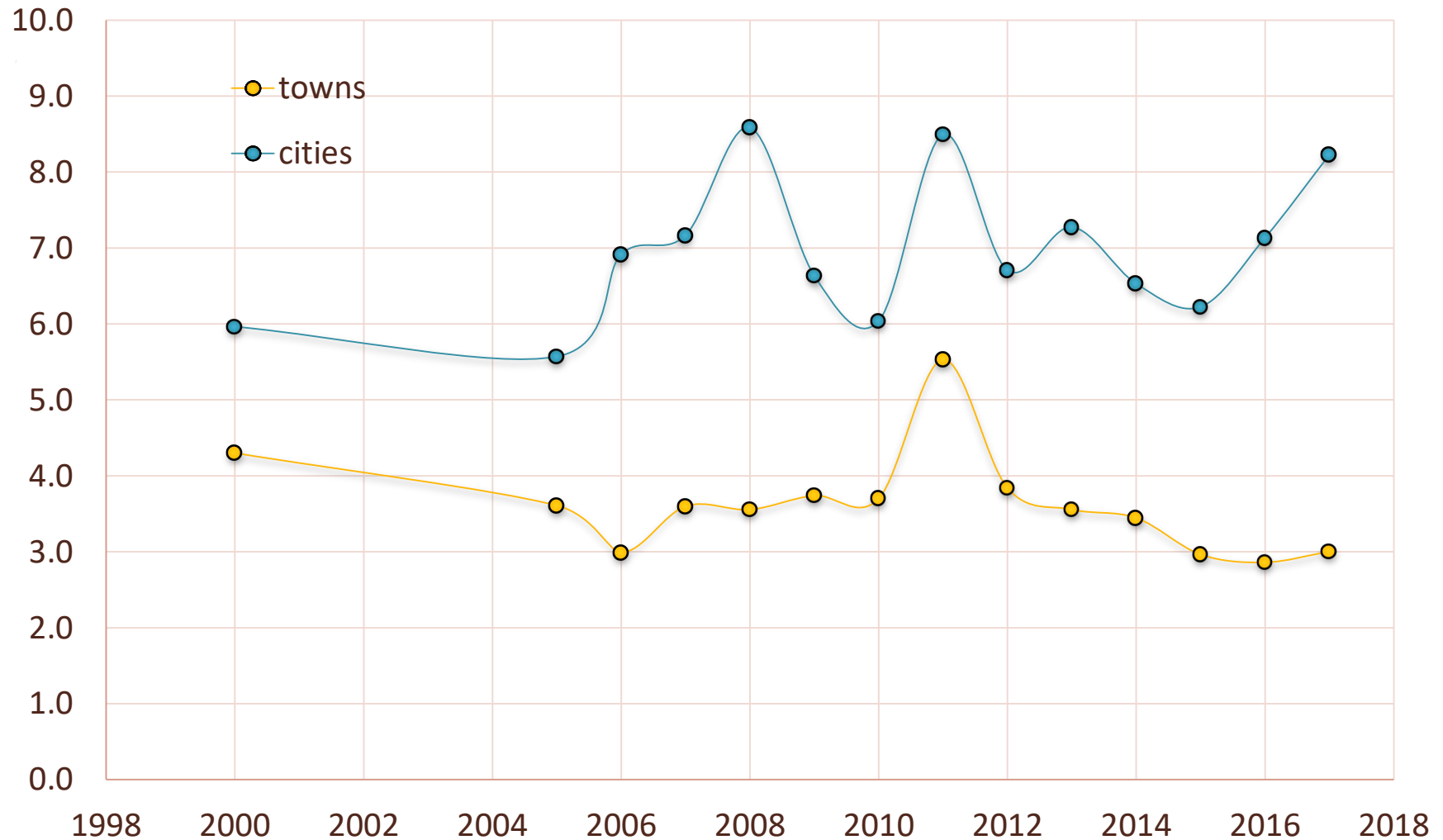


Mortality trends by rural/urban (unadjusted e_0^t)



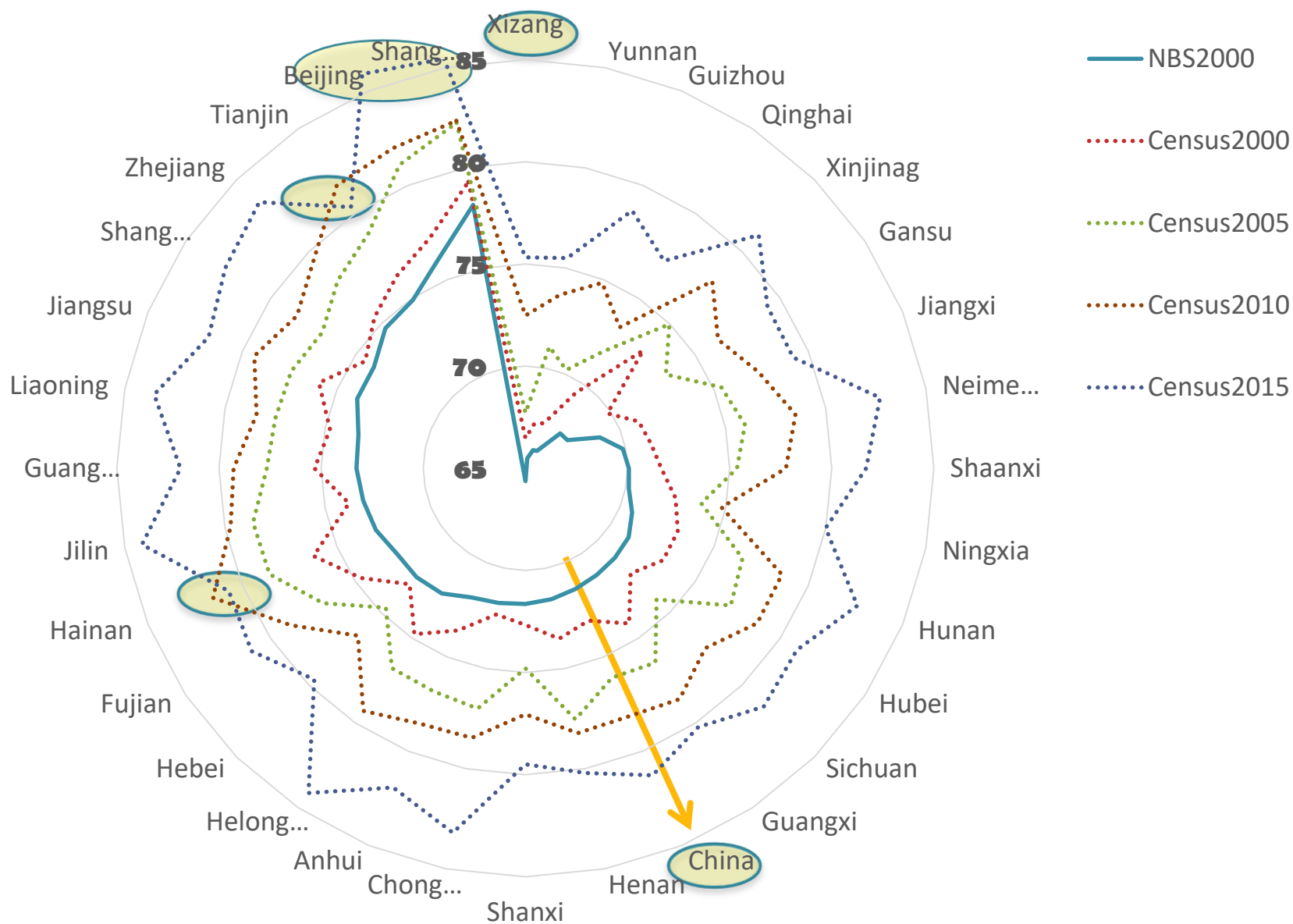


Mortality trends: rural/urban gap (unadjusted e_0^t) comparing to e_0^t for villages



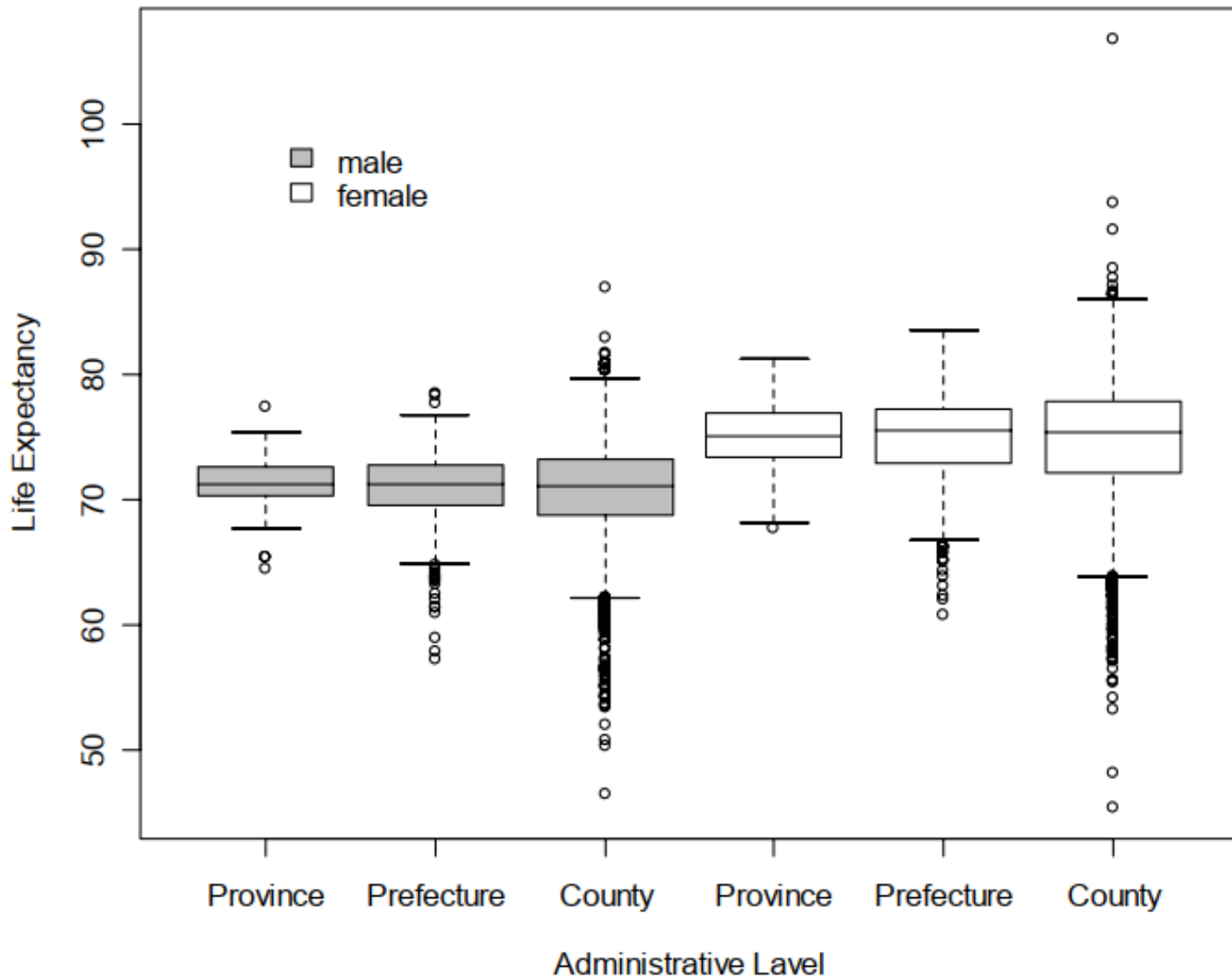


Regional trends and variation, e_0^t





Mortality variation: e_0 (2000, unadjusted)





Mortality variation: county *level* e_0 (2000, raw)

Variable	Mean	S.D.	Percentile				
			Min	5	Med	95	Max
e_0^m	69.9	3.9	46.4	62.9	70.5	75.3	86.9
e_0^f	73.8	4.7	45.3	65.4	74.5	80.0	106.7
${}_1q_0^m$	25.7	23.4	0.0	4.6	22.9	91.1	376.7
${}_1q_0^f$	32.8	32.0	0.0	4.9	18.9	68.6	307.3



Brass Logit Model Smoothing

- **Brass (1971) relational model: linear relationship between l_x 's of an observed population and of a standard population at logit scale**

$$\text{logit}(l_x) = \alpha + \beta \cdot \text{logit}(l_x^s)$$

- **Consistency of mortality across different age groups**
- **Choice of standard life table: “must be some kind of average Brass (1971)”. We use provincial average.**



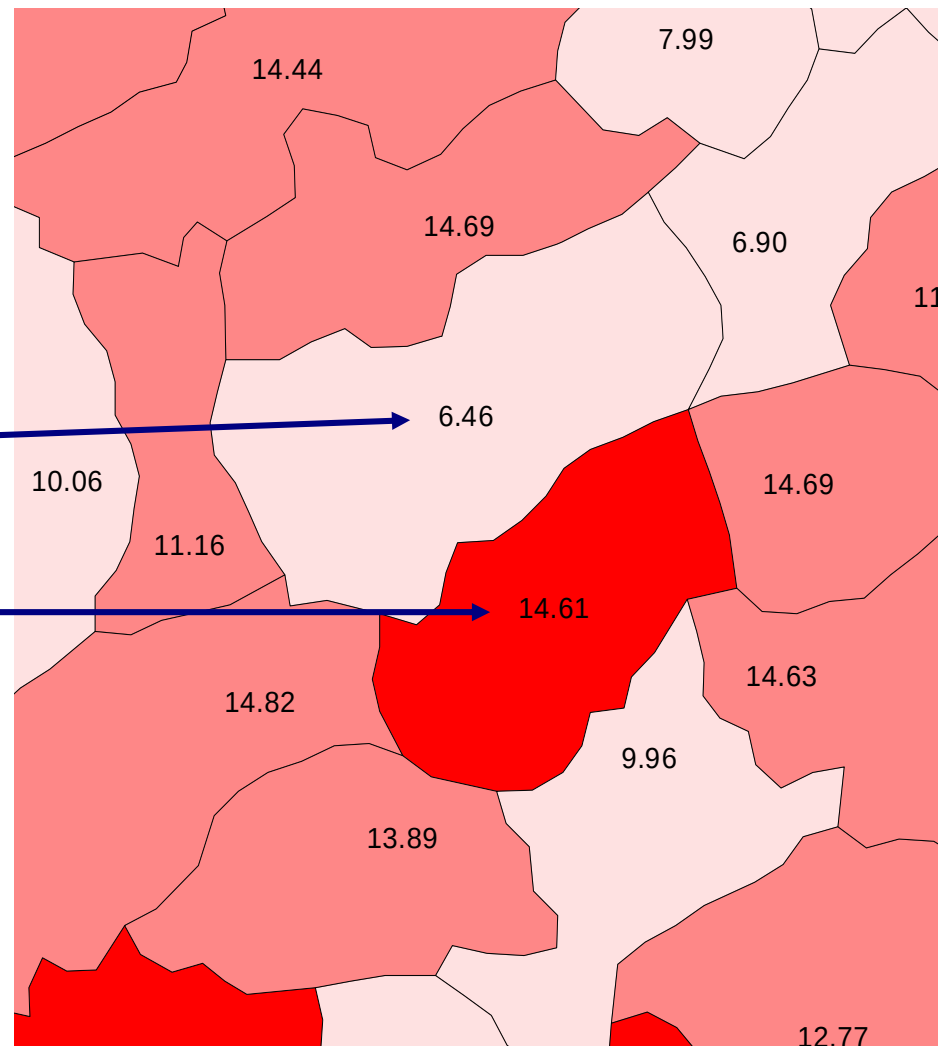
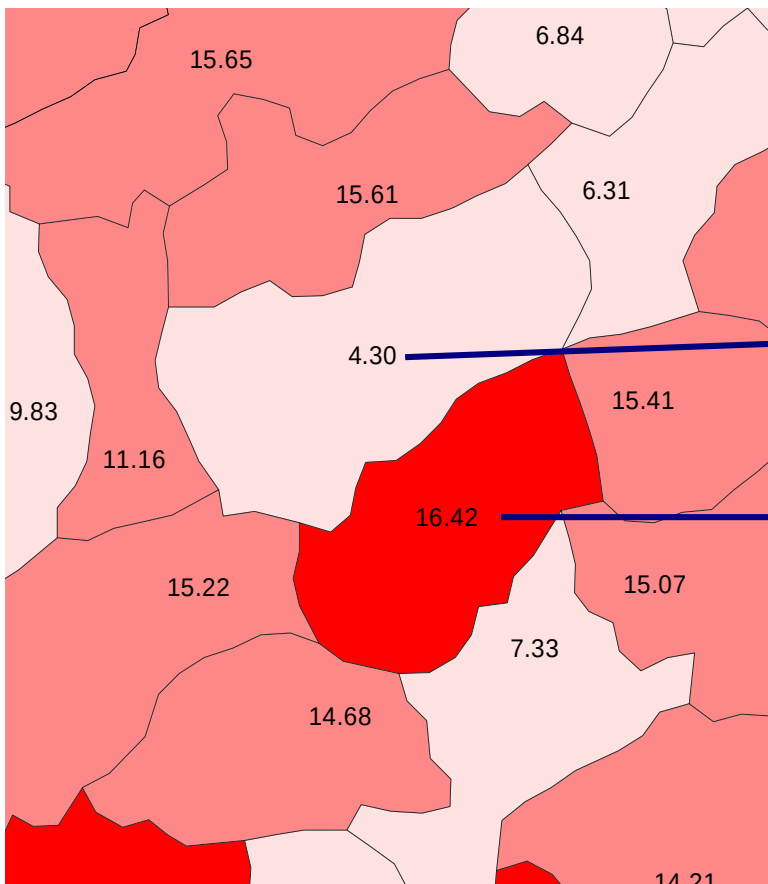
Empirical Bayes Estimates

- Empirical Bayes Estimates (Efron and Morris 1973, 1975; Clayton and Kaldor 1987; Marshall 1991): pool information across areas to produce more stable and robust estimates using the Empirical Bayes methods.
- Information from neighboring units serves as *priori* for the local estimate.

$$\hat{\theta}_i = m_n + \frac{S^2 - \frac{m_n}{\bar{n}}}{S^2 - \frac{m_n}{\bar{n}} + \frac{m_n}{n_i}} (m_i - m_n)$$



Empirical Bayes Estimates



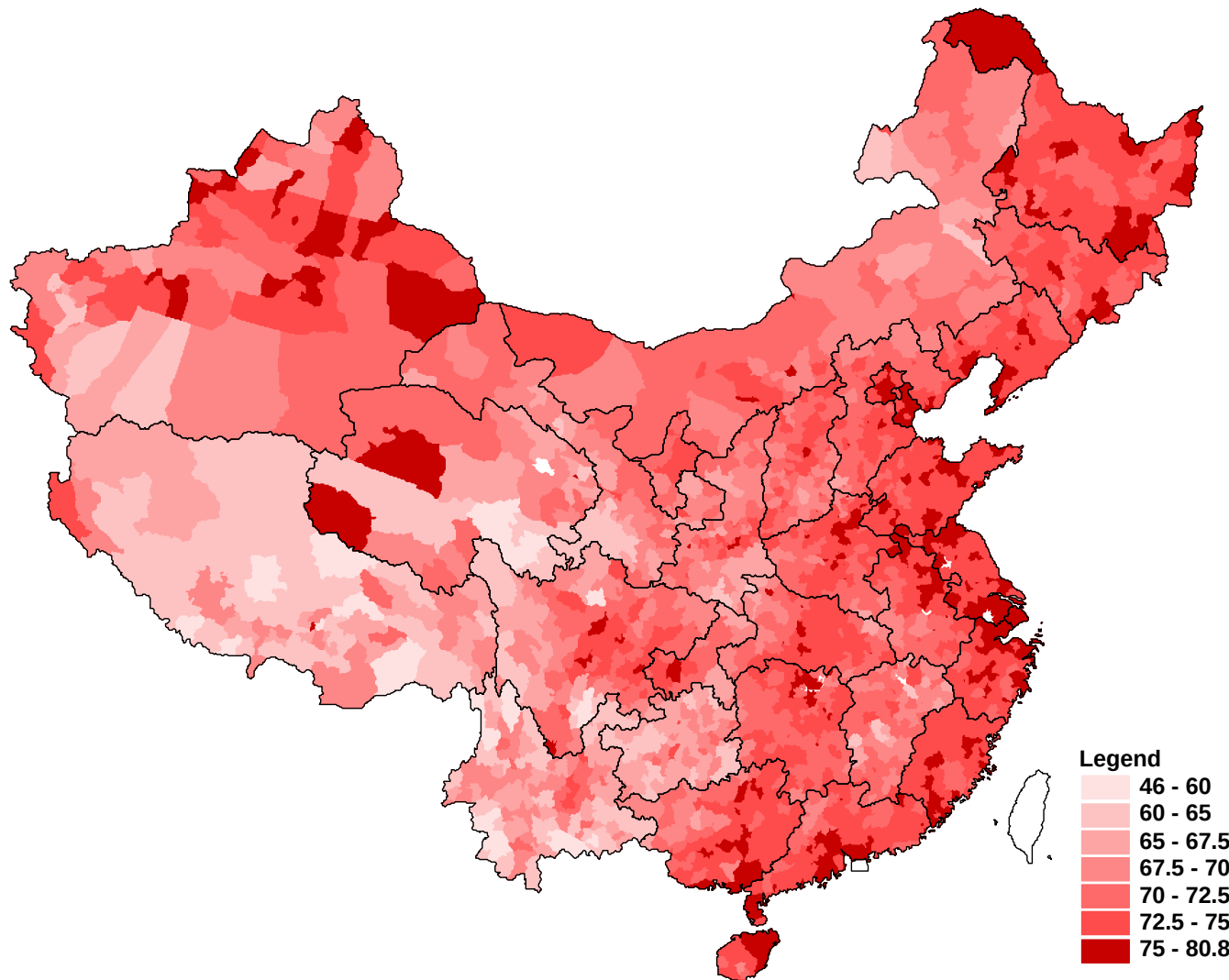


Mortality variation: county level e_0 (2000 smoothed)

Variable	Mean	S.D.	Percentile				
			Min	5	Med	95	Max
e_0^m	69.8	3.7	46.6	62.9	70.5	74.5	79.3
e_0^f	73.4	4.2	45.5	65.3	74.3	78.6	84.0
${}_1q_0^m$	26.0	22.1	0.9	6.1	23.8	93.0	355.2
${}_1q_0^f$	33.6	31.3	1.2	6.6	19.2	66.9	267.7

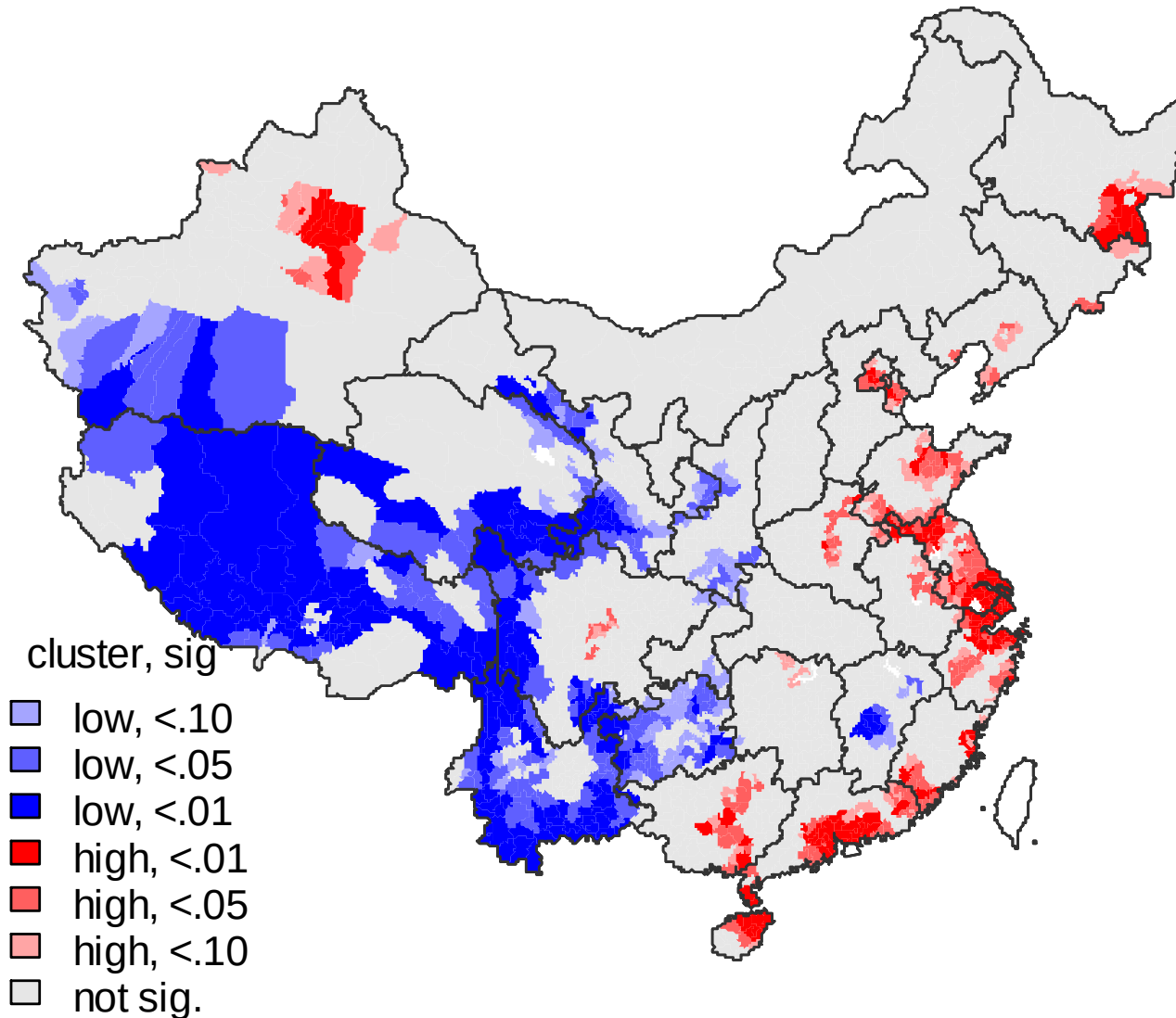


Mortality variation: county level e_0^t (2000)





Mortality variation: county level e_o^f clusters (2000)



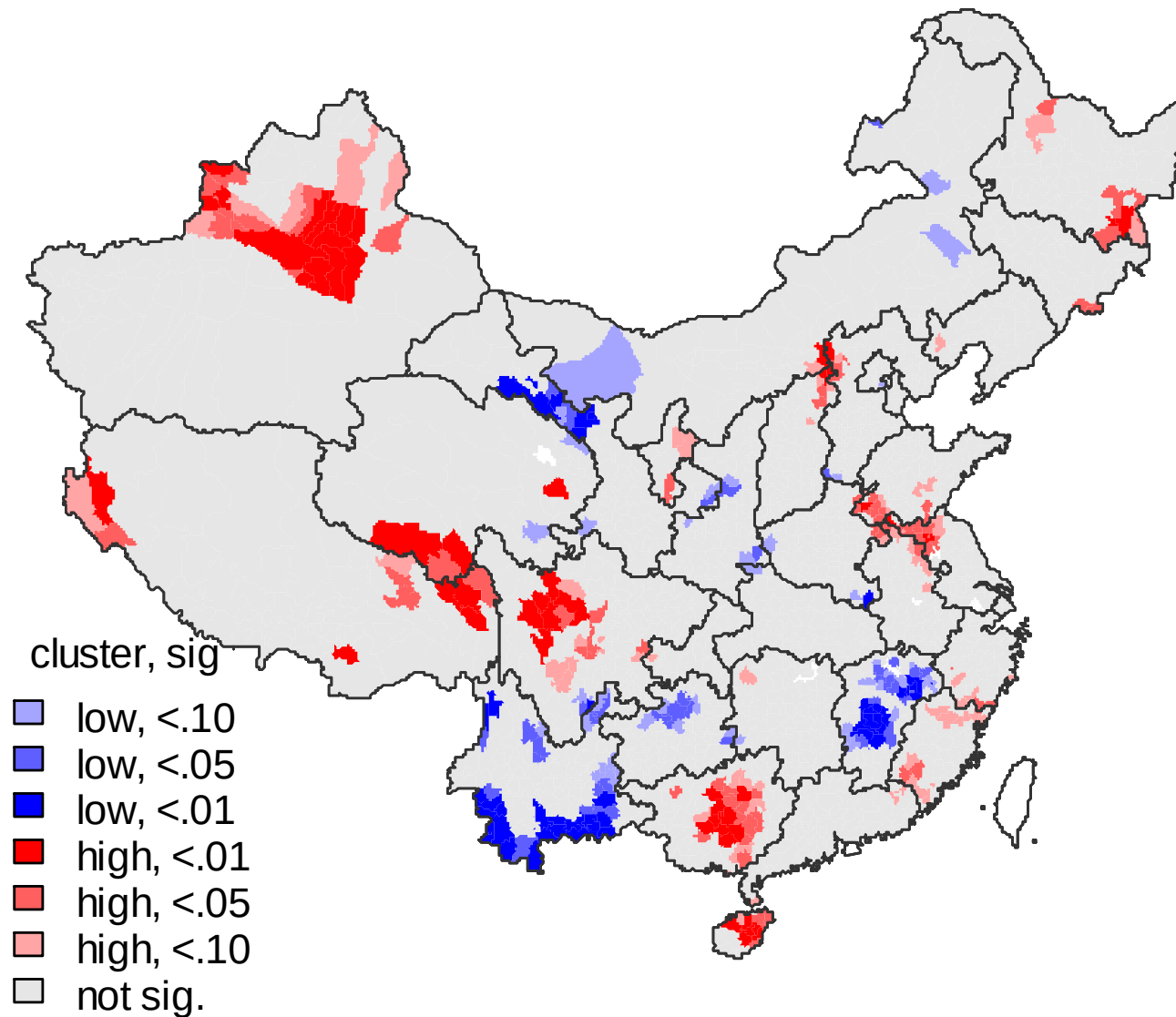


Mortality variation: e_0 2000 regression

Predictor	Male Model				Female Model				Difference	
	coef	std	beta		coef	std	beta		(Male - Female)	
(Intercept)	61.526	1.9371		***	64.099	2.3035		***	-2.573	
GDP	0.6292	0.1119	0.1205	***	1.0133	0.1331	0.1676	***	-0.384	*
Urbanization	0.0046	0.0042	0.0296		-0.0014	0.0050	-0.0080		0.006	
Employment	0.0179	0.0089	0.0435	*	0.0312	0.0106	0.0654	**	-0.013	
Health Care	-0.6532	0.2782	-0.0518	*	-0.4184	0.3308	-0.0286		-0.235	
Water	0.0044	0.0026	0.0337		0.0040	0.0031	0.0265		0.000	
Lavatory	0.0058	0.0022	0.0448	**	0.0069	0.0026	0.0458	**	-0.001	
Bath Facility	0.0121	0.0042	0.0659	**	0.0162	0.0050	0.0761	**	-0.004	
Education	0.9515	0.0982	0.3524	***	1.0719	0.1167	0.3428	***	-0.120	
Discrimination	0.0465	0.1629	0.0060		0.0353	0.1937	0.0039		0.011	
Ethnicity	-0.0253	0.0029	-0.2152	***	-0.0151	0.0035	-0.1107	***	-0.010	*
Marriage	-0.0417	0.0133	-0.0677	**	-0.0705	0.0158	-0.0987	***	0.029	
Family Structure	0.0117	0.0095	0.0229		0.0106	0.0113	0.0180		0.001	
Latitude	0.0324	0.0177	0.0580		-0.0023	0.0210	-0.0035		0.035	
Longitude	0.0064	0.0153	0.0167		0.0168	0.0182	0.0379		-0.010	
Elevation	-0.0003	0.0001	-0.0780		-0.0002	0.0002	-0.0415		0.000	
Terrain	-0.0009	0.0003	-0.0544	**	-0.0011	0.0004	-0.0548	**	0.000	
*** denotes p<.001; ** denotes p< .01; * denotes p< .05										



Mortality variation: county level e_o^f (2000)





Mortality variation: observations

- **Huge regional variation: China is a world of its own!**
- **Regional variation is relatively stable**
- **Regional differences are largely driven by developmental factors**



Next Steps

- **A model/system simultaneously estimating fertility, mortality, population age structure using all available survey/census data to consistently estimate demographic parameters.**
- **Subnational evaluation**
- **HMD standards/protocols**



Thank you!