Evaluation of the Completeness of Birth Registration in China Using Analytical Methods and Multiple Sources of Data

Wei Chen
Center for Population and Development Studies
Renmin University of China
Email: weichen@ruc.edu.cn

Background

- China's below-replacement fertility and deterioration of the quality of demographic data
- Recent level of fertility in China is a matter of considerable uncertainty and controversy
- Births and child population have long been under-reported in all population surveys
- No consensus on the completeness of birth registration and fertility level in China

China's birth registration

- Household registration system (Hukou) by Ministry of Public Security
- Family planning statistics by National Population and Family Planning Commission
- Hospitalized delivery statistics by Ministry of Health
- Population census and annual population survey by National Bureau of Statistics

Prior Studies

- Since birth registration data are rarely used in demographic analysis, virtually no studies have evaluated the birth data from the various government departments
- Few studies assessed the completeness of birth report in the population censuses

Prior Studies

- Using primary school enrollment data, births were under reported by 18% in 2000 census (Zhang and Cui 2003)
- Assuming a pattern of under-reporting by age similar to 1990 census, 28% of births in 2000 census were not reported (Wang 2003)
- Assessed by the NBS-published number of births, 2010 census had 15% of births not recorded (Cui et al. 2013)
- According to the average population at age 0 from 2010 census and estimated deaths at age 0, 15.4% of births were under reported in the 2010 census

Purpose of this study

- This research brings together multiple sources of data and different methods to reach a consistent estimation of the completeness of China's birth registration over 2000-2010
  - Estimating annual births over 2000-2009
  - Estimating births in 2010, and
  - Estimating average inter-censal births to be validation of the above results
  - Evaluating the completeness of birth registration by comparing births from the various government departments to the above results

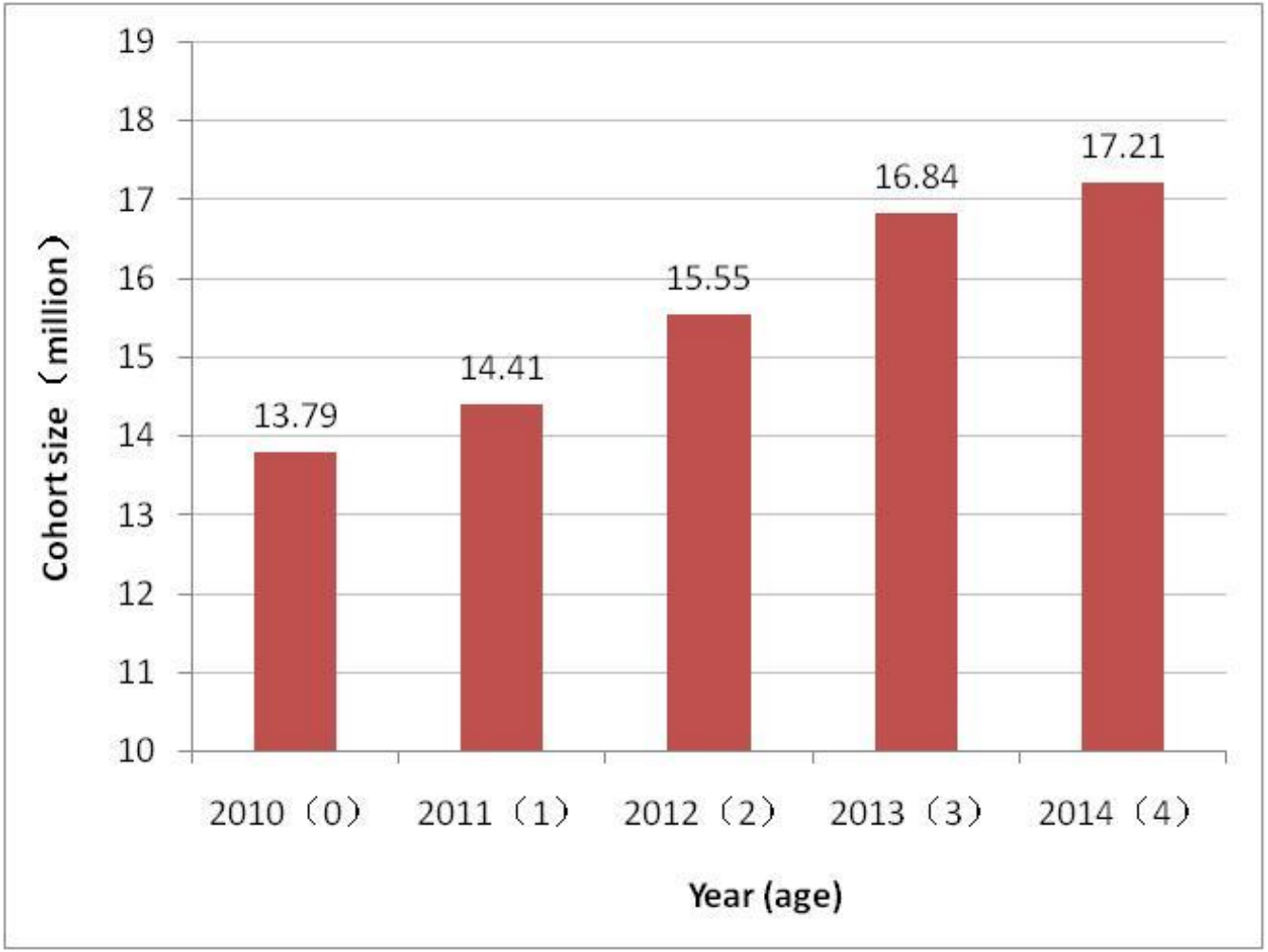
Data and Methods

- Using data from census, education enrolment and household registration to estimate annual births over 2000-2009
- Using an indirect estimation method (P/F ratio) to estimate births in 2010
- Using another indirect method (generalized stable population model) to estimate average births in inter-censal periods
- Evaluating the completeness of birth registration data from the various government departments by comparing to the above results

Under-reporting in population censuses

Age	2000 census	Age	2010 census	Education data	Hukou data
（1）	（2）	（3）	（4）	（5）	（6）
0	1379	**10**	1445	1636	1479
1	1150	**11**	1394	1614	1443
2	1401	**12**	1540	1692	1546
3	1445	**13**	1523	1737	1562
4	1522	**14**	1589	1791	1635
5	1693	**15**	1802	1849	1744
6	1647	**16**	1879	1878	1743
7	1791	**17**	2078	1968	1850
8	1875	**18**	2076	2049	1924
9	2008	**19**	2154	2204	2057

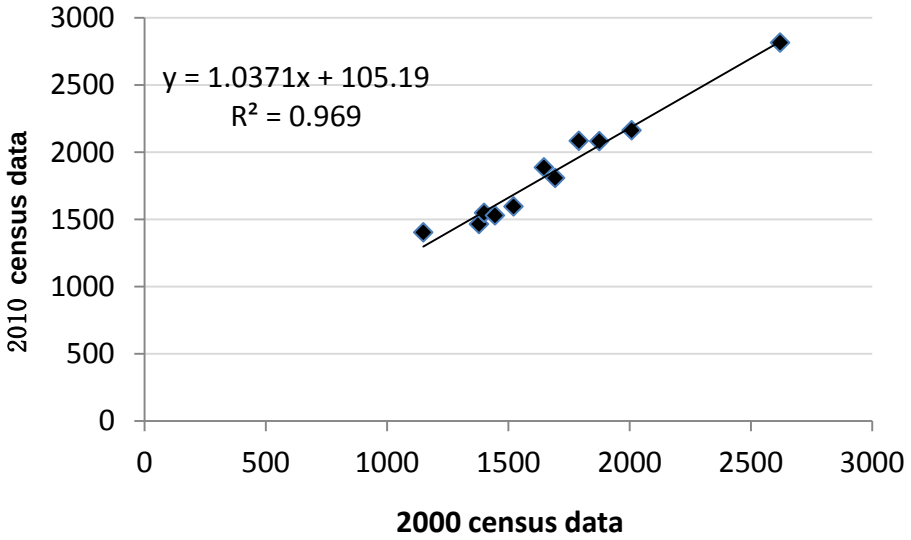
Population at age 0 in 2010 census, counted in 2011-2014 annual population surveys

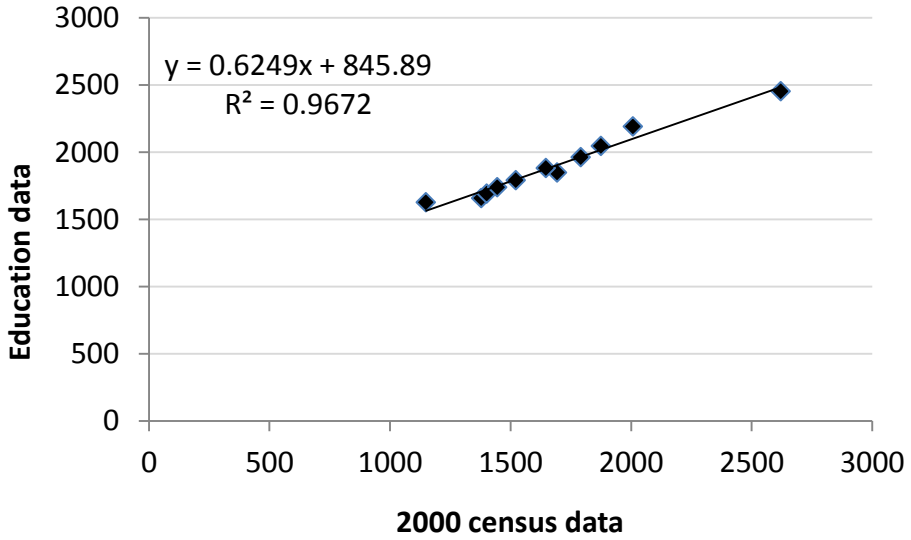


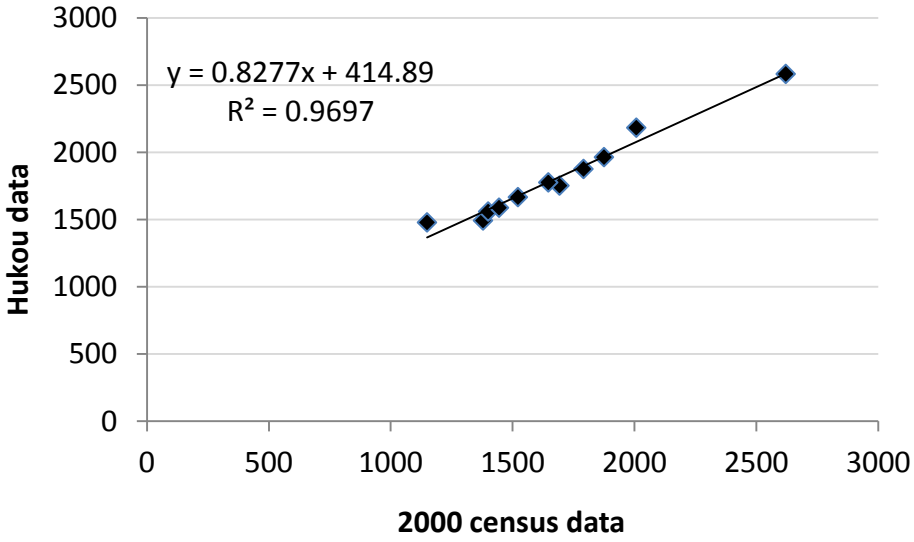
Estimating Annual Number of Births over 2000-2009

- Using life table survival ratios to backward project 2000 age 0-10 population from 2010 census age data 10-20, education data and Hukou data
- Calculating linear regression models of reported 2000 census age data 0-10 against the backward projected 2000 age data 0-10

Linear regression of reported 2000 census data against the backward projected 2000 data







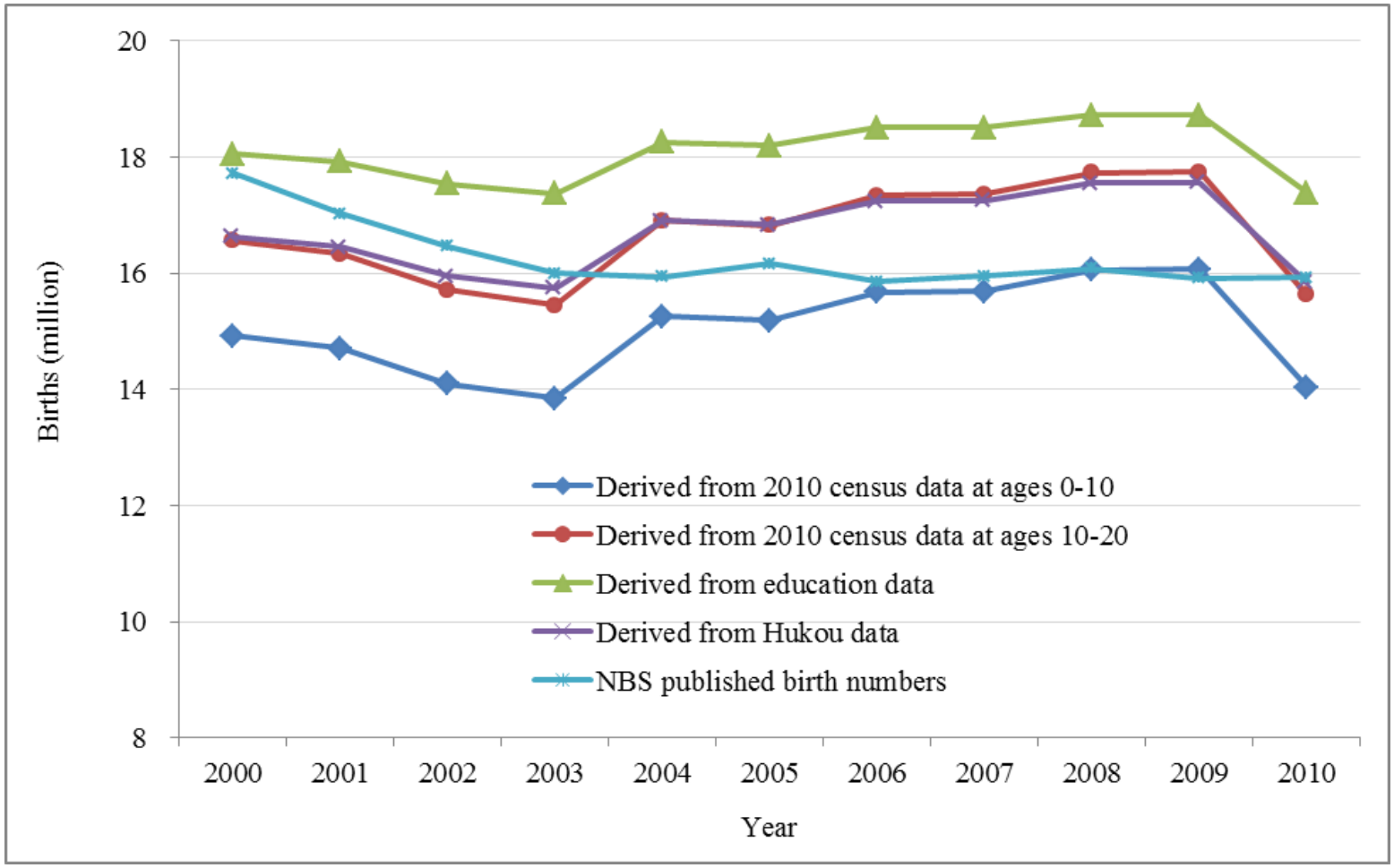
Annual births in 2000-2009

- Plugging into the regression equations the reported 2010 census age data 0-10 to get the adjusted 2010 census age data 0-10
- Using life table survival ratios to backward project annual number of births over 2000-2010 from the estimated 2010 age data 0-10

Annual births in 2000-2010

Source of data	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Derived from 2010 census data at ages 0-10	14.92	14.70	14.10	13.85	15.26	15.18	15.67	15.69	16.05	16.07	14.03
Derived from 2010 census data at ages 10-20	16.56	16.33	15.70	15.44	16.91	16.82	17.33	17.35	17.73	17.74	15.63
Derived from education data	18.05	17.92	17.53	17.37	18.25	18.20	18.50	18.51	18.72	18.72	17.38
Derived from Hukou data	16.63	16.45	15.95	15.74	16.91	16.84	17.24	17.25	17.55	17.56	15.84
NBS published birth numbers	17.71	17.02	16.47	15.99	15.93	16.17	15.85	15.95	16.08	15.91	15.92

Annual births in 2000-2010



Reestimating the 2010 births

- Using fertility data from the 2010 census and the P/F ratio method (Brass, 1967; United Nations, 1983)
- Fertility data: age-specific fertility rates and average number of child ever born by age of mother

P/F ratio method to estimating births

i	Age	$P(i)$	$f(i)$	$\phi(i)$	$F(i)$	P/F Ratio
1	15-19	0.0125	0.0059	0.0296	-0.0002	
2	20-24	0.2596	0.0695	0.3770	0.1822	1.4252
3	25-29	0.8440	0.0841	0.7973	0.5975	1.4126
4	30-34	1.2928	0.0458	1.0265	0.9249	1.3978
5	35-39	1.5235	0.0187	1.1201	1.0777	1.4136
6	40-44	1.6871	0.0075	1.1576	1.1318	1.4907
7	45-49	1.8366	0.0047	1.1810	1.1739	1.5645

$$F(i) = \phi(i-1) + a(i)\,f(i) + b(i)\,f(i+1) + c(i)\phi(7)$$

Coefficients for converting $\phi(i)$ into $F(i)$

i	Age	*a*	*b*	*c*
1	15-19	2.147	-0.244	0.0034
2	20-24	2.838	-0.758	0.0162
3	25-29	2.760	-0.594	0.0133
4	30-34	2.949	-0.566	0.0025
5	35-39	3.029	-0.823	0.0006
6	40-44	3.419	-2.966	-0.0001
7	45-49	3.535	-0.007	-0.0002

Number of births in 2010

- P/F ratios across 20-24 to 35-39 is roughly 1.4, implying that the period fertility in 2010 census is under-estimated by 40%
- Adjusting the births from the 2010 census: 11.9 million is multiplied by 1.41, resulting an adjusted number of births of 16.78 million, implying a TFR of 1.66

Estimating the average annual number of births over 2000-2010

- Using age distribution data over age 10 from two censuses and generalized stable population models to estimate the average number of births between the two censuses
- The integrated estimation method developed by Preston (1983)

The integrated approach

$$\frac{e^{-\int_0^a r(x)dx}}{c(a)} = \frac{1}{b} + \frac{k}{b}\frac{q_s(a)}{p_s(a)}$$

- *b* is an estimate of the crude birth rate, and *k* adjusts the standard life table mortality curve to the level of mortality in the data of the population under study
- The life tables of the 1982 census as the standard

Chinese female, 2000-2010

Age	$N(a)_{2000}$	$N(a)_{2010}$	$N(a)$	$r(a)$	$\int_0^a r(x)dx$	$c(a)$	$\frac{e^{-\int_0^a r(x)dx}}{c(a)}$	$\frac{q_s(a)}{p_s(a)}$
0	31329680	34470044	32874867	0.010	0.000	---	---	---
5	41849379	32416884	36932596	-0.026	0.010	0.011	84.868	0.0540784
10	60051894	34641185	46187366	-0.055	-0.016	0.013	80.984	0.0597694
15	50152995	47984284	49060651	-0.004	-0.071	0.015	93.050	0.0632462
20	46635408	63403945	54591123	0.031	-0.075	0.017	87.417	0.0682241
25	57371507	50176814	53693847	-0.013	-0.045	0.017	71.764	0.0753862
30	61953842	47616381	54470990	-0.026	-0.058	0.017	76.822	0.0833816
35	53005904	57634855	55288087	0.008	-0.084	0.018	86.354	0.0925701
40	38999758	61145286	49245407	0.045	-0.076	0.017	86.954	0.1044763
45	41581442	51818135	46512194	0.022	-0.031	0.015	75.809	0.1206473
50	30500075	38389937	34293873	0.023	-0.009	0.013	80.475	0.1445921
55	22308869	40229536	30393735	0.059	0.014	0.010	89.603	0.1831633
60	20029370	28832856	24164432	0.036	0.073	0.009	79.114	0.2453695
65	17231112	20364811	18754347	0.017	0.109	0.007	83.821	0.3576617
70	13137995	16568944	14787191	0.023	0.126	0.005	98.658	0.5557204
75	8752519	12573274	10547815	0.036	0.149	0.004	116.309	0.9623856
80	4785290	7455696	6022136	0.044	0.185	0.003	148.373	1.8168888
85+	2659218	4758898	3607797	0.058	0.230	0.002	204.533	

The female equation

- $Y = 76.396 + 44.149X$, implying

$$b = \frac{1}{76.396} = 0.013090$$

$$k = \frac{44.149}{76.396} = 0.57790$$

Chinese male, 2000-2010

Age	$N(a)_{2000}$	$N(a)_{2010}$	$N(a)$	$r(a)$	$\int_0^a r(x)dx$	$c(a)$	$\frac{e^{-\int_0^a r(x)dx}}{c(a)}$	$\frac{q_s(a)}{p_s(a)}$
0	37648694	41062566	39330940	0.009	0.000	---	---	---
5	48303208	38464665	43197364	-0.023	0.009	0.013	76.188	0.0542455
10	65344739	40267277	51798193	-0.048	-0.014	0.014	74.172	0.0611322
15	52878170	51904830	52389993	-0.002	-0.063	0.016	86.150	0.0653082
20	47937766	64008573	55586517	0.029	-0.064	0.016	83.903	0.0710921
25	60230758	50837038	55401230	-0.017	-0.035	0.017	70.640	0.0787078
30	65360456	49521822	57075335	-0.028	-0.052	0.017	75.873	0.0866184
35	56141391	60391104	58240409	0.007	-0.080	0.018	85.019	0.0962545
40	42243187	63608678	52199206	0.041	-0.073	0.017	85.593	0.1097041
45	43939603	53776418	48692521	0.020	-0.032	0.015	76.353	0.1293549
50	32804125	40363234	36453148	0.021	-0.012	0.013	81.781	0.1600736
55	24061506	41082938	31816974	0.053	0.009	0.010	91.951	0.2122488
60	21674478	29834426	25537543	0.032	0.063	0.009	83.762	0.3032271
65	17549348	20748471	19104288	0.017	0.094	0.007	91.725	0.4747560
70	12436154	16403453	14328380	0.028	0.111	0.005	112.641	0.7918835
75	7175811	11278859	9073238	0.045	0.139	0.004	140.120	1.4776122
80	3203868	5917502	4422801	0.061	0.184	0.002	193.794	3.0639912
85+	1342707	2857250	2005561	0.076	0.245	0.001	299.375	

The male equation

- $Y = 75.217 + 45.326X$, implying

$$b = \frac{1}{75.217} = 0.013295$$

$$k = \frac{45.326}{75.217} = 0.60261$$

Inter-censal fertility derived from the integrated approach (comparison with NBS published estimates)

Fertility measure	1982-1990	1990-2000	2000-2010
Estimated crude birth rate	22.33	15.08	13.18
Published crude birth rate	21.58	17.25	12.50
Estimated births（million）	**23.78**	**17.80**	**16.88**
Published births（million）	23.09	20.67	16.27
Estimated total fertility rate	2.63	1.68	1.56

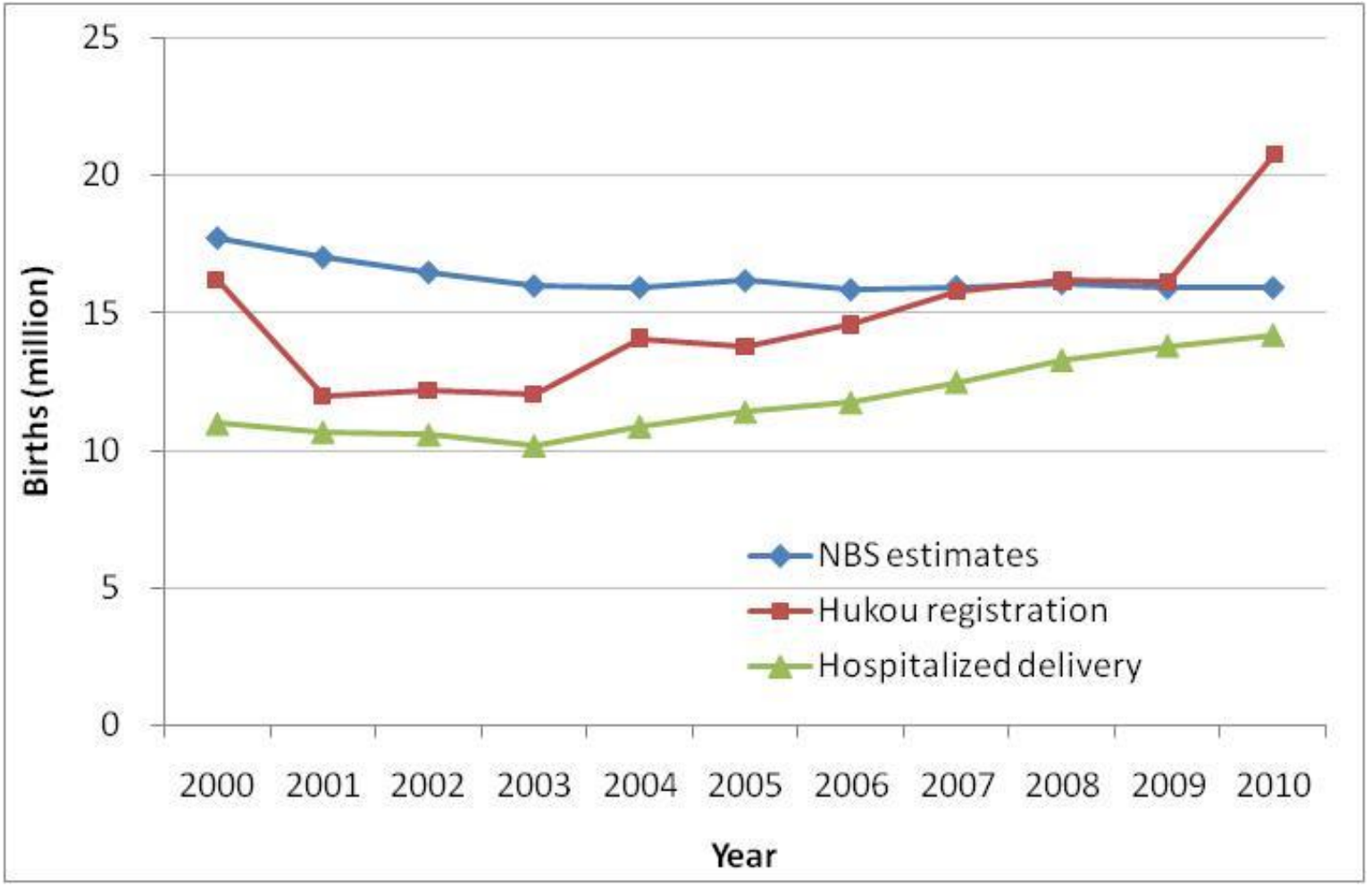
Estimated average number of births over 2000-2009 from the regression method

Sources of data	Row	Average over 2000-2009
Derived from 2010 census data at ages 0-10	(1)	15.15
Derived from 2010 census data at ages 10-20	(2)	16.79
Derived from education data	(3)	18.18
Derived from Hukou data	(4)	16.81
NBS published birth numbers	(5)	16.31
Average	(6)	16.65

Birth registration 2000-2010: various sources (million)

Year	NBS estimates	Hukou registration	Hospitalized delivery
2000	17.71	16.21	10.99
2001	17.02	11.96	10.69
2002	16.47	12.16	10.59
2003	15.99	12.01	10.19
2004	15.93	14.07	10.89
2005	16.17	13.74	11.42
2006	15.85	14.56	11.77
2007	15.95	15.77	12.51
2008	16.08	16.16	13.31
2009	15.91	16.15	13.83
2010	15.92	20.78	14.22
Average	**16.27**	**14.87**	**11.85**

Birth registration 2000-2010: various sources



Completeness of birth registration

- Judging by the result from the Preston integrated approach
  - The NBS estimates are 96% complete
  - Hukou registration data are 88% complete
  - Hospitalized delivery data are 70% complete
- Birth registration data are more complete in the late 2000s than in the early 2000s

Conclusion and Discussion

- Different data and different methods have arrived largely consistent results
  - The average number of births over 2000-2010 stands at around 16.6-16.8 million
  - The average of the NBS-published annual estimates is only slightly lower than our estimate
- Referring to the results from this study, health department recorded nearly 90% while Hukou registration department only 70% of the births over 2000-2010

Conclusion and Discussion

- The integrated approach developed by Preston (1983) is more promising and of much higher application value than other two methods
  - Population stability is a no more needed assumption, so is the assumption of population closed to migration
  - Only population age distributions at two censuses are needed to estimate fertility and mortality
  - Preston (1983) performed analysis of sensitivity to various forms of error, and it appears to provide a fairly robust estimate of fertility
  - The relative coverage completeness of the two censuses rather than the absolute coverage completeness matters

Conclusion and Discussion

- Birth registration in China has been apparently affected by the one-child policy
  - Millions of births or children, more female than male, are deliberately under-reported in population censuses or surveys
  - Fertility policy is more important than estimation method for achieving accurate birth or population data

Conclusion and Discussion

- China adjusted the one-child policy and started to implement a two-child policy in 2016
  - No more hiding is needed for second births
  - The two-child policy is hoped to help improve quality of birth or population data in China

Thanks