Health, Education and Income Trends in the US
An paper in Honor of Claudia Goldin

Hoyt Bleakley

Dora Costa

Adriana Lleras-Muney

Very long run stylized facts

- Health, education and wages remained stagnant for most of humanity
- They have grown very rapidly in the last few centuries in the world as a whole

- This paper: US in 19th and 20th century exploration

Claudia's work on education and income in the US

- Carefully and painfully documented the trends and patterns in educational achievement:
  - 1900-1950 HS graduation rates: 9 to 50%
  - 1950-2000 slowdown
  - Examined factors driving US early leadership

- Documented U shape trends in income inequality:
  - Declining up to 1950
  - Increasing thereafter

- Documented secular changes in returns to schooling (U shape)

19th and 20th century health

- 19th century – ups and downs
- 20th century just as equally large and unprecedented changes in health:
  - LE increased about 30 years
  - 1900-1950: IMR and infectious diseases
 - 1950 onwards: adult MR and chronic diseases

This paper

- New evidence on 20th century trends in early health

  - Rise in birth weights
  - Improvements in mothers' health
  - Decline in within sibling variance

- What is the contribution of health improvements to education and income?

  - Cannot assess causality but document correlations between measures of early health and educational attainment over time

Tentative conclusions

- Strong direct effects of health throughout 19th and 20th C
  - Height and income, wealth, mortality

- Later emergence/strengthening of "skill bias" of health
  - Height predicts education weakly 19th C and pre-WWII
  - Height predicts education strongly latter ½ or 20th C

- Changing importance of brain vs brawn, variance of "insults"
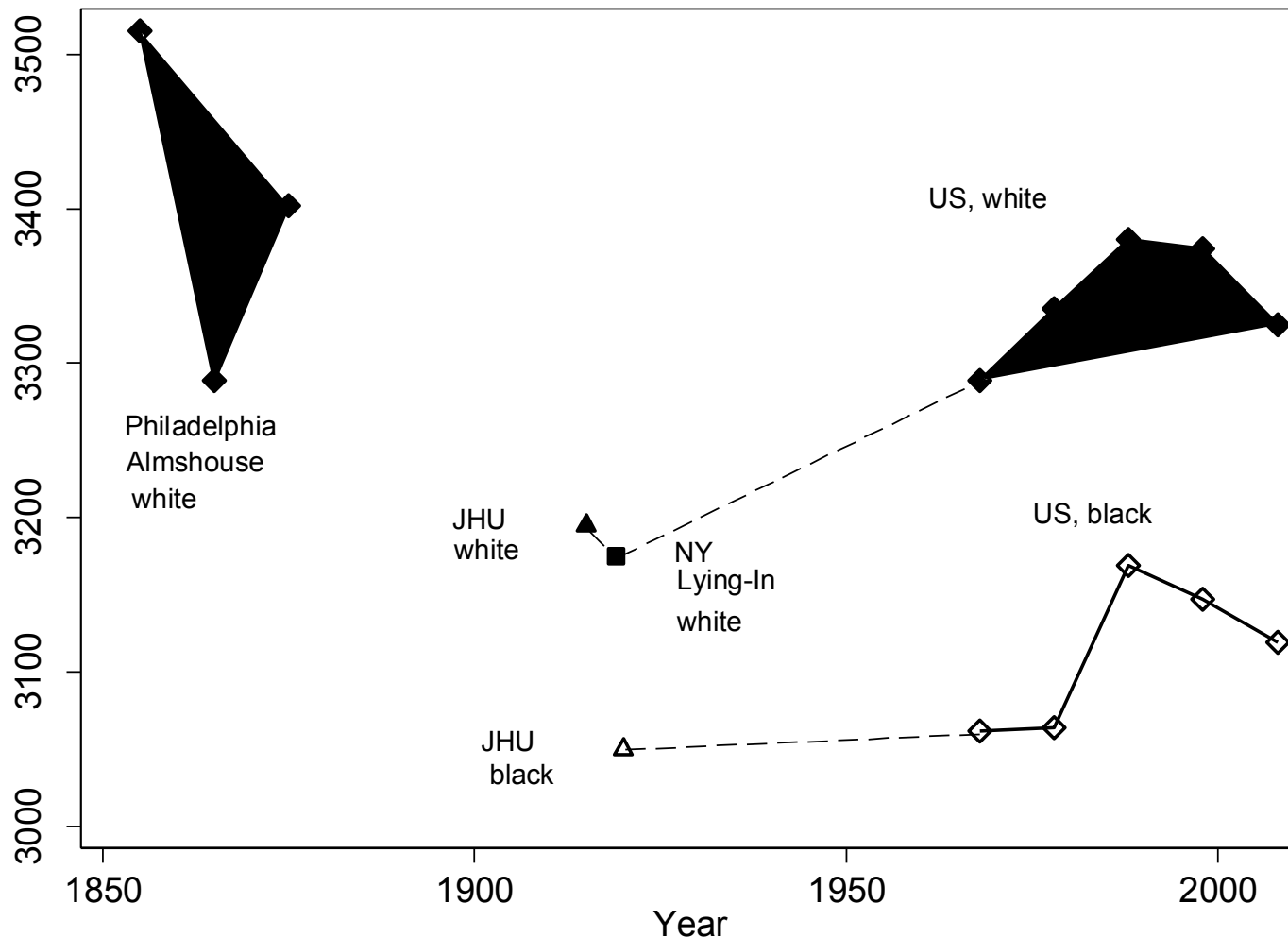
Know a lot about long-run height trends …
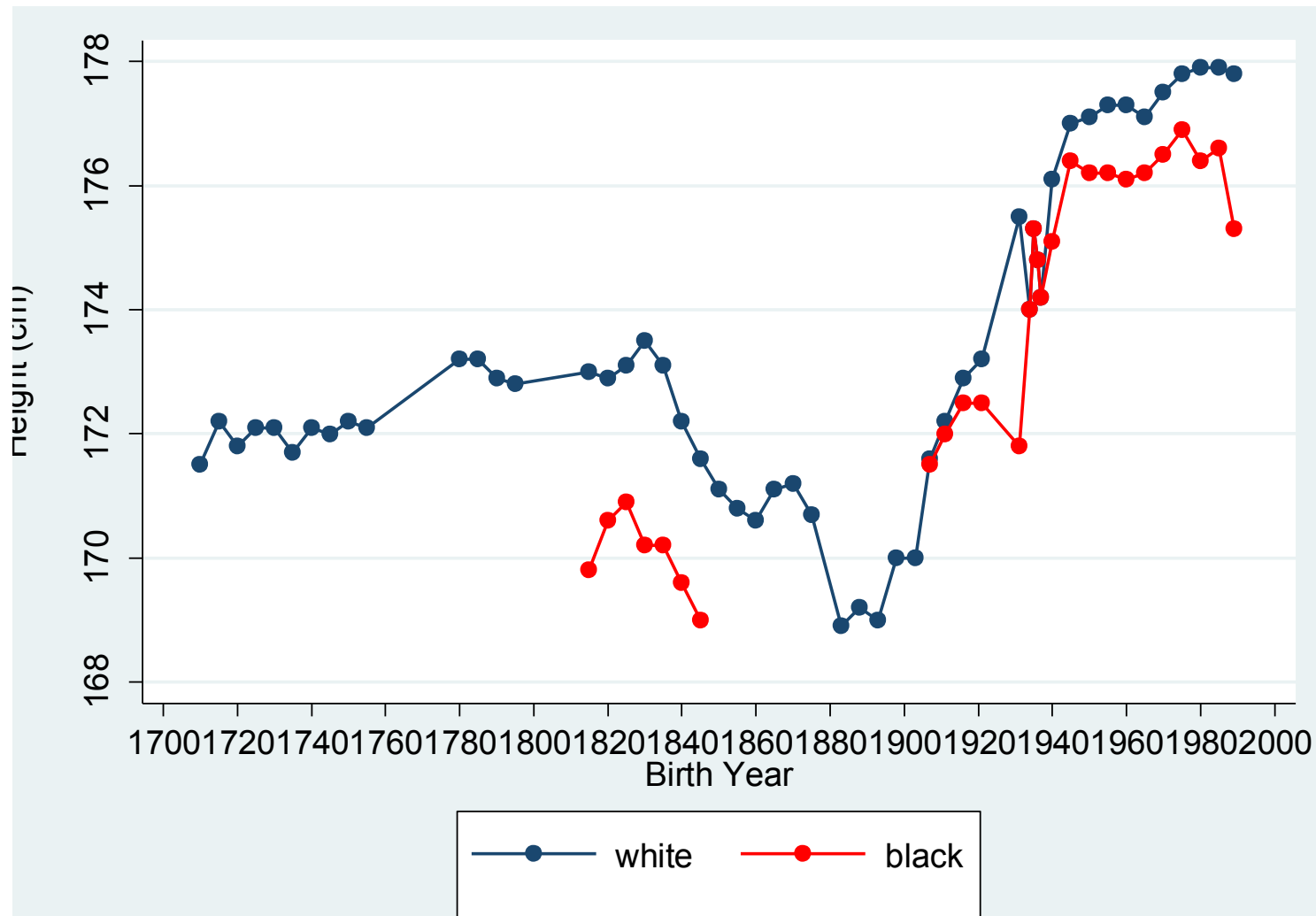
Puzzle about birth weights ...
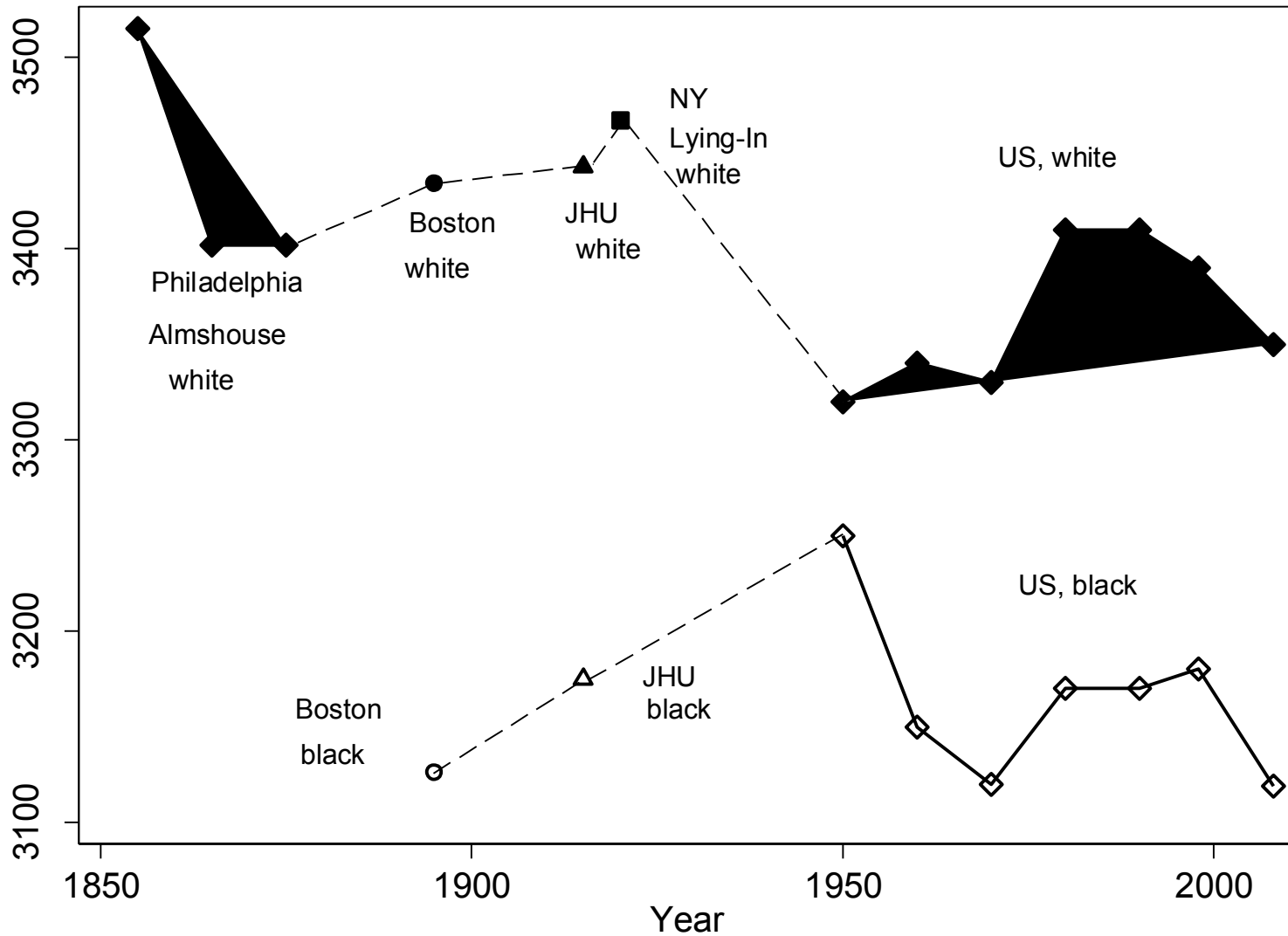
Consistency if look at first births; but why higher order births so heavy?

Mothers' health has improved

	NY Lying-In	JHU White	JHU Black	1988 White	1988 black
Height (cm)	157.0	159.4	160.8	164.1	164.2
At least one BLP reading High Blood Pressure Systolic > 140 or Diastolic > 90		31.8	29.8	15.5	12.2
Severely High Blood Pressure Systolic > 160 or Diastolic > 105		10.0	6.4	3.2	2.6
Syphilis		2.0	13.4	0.1	1.5

Increase in sibling correlations (Whites only)…

	Adult Height, Brother-Brother	Birthweight full term, singletons	Birth Length Full term, singletons	Gestational Wks (full term, singletons)
Union Army, 1861-5	0.394 (0.024)			
WWII, 1939-45	0.462 (0.024)			
Norway ,1984-86	0.47			
Mazumder, PSID 1951-1968 cohort	0.492 (0.017)			
JHU 1900-35		0.466 (0.061)	0.213 (0.075)	0.124 (0.068)
Norway, 1967-2004		0.506 (0.003)	0.408 (0.004)	0.316 (0.004)
Mazumder, PSID 1985-1997 cohort		0.500 (0.019)		0.377 (0.019)

1-health and education

1. Physical health: height

2. cognitive ability/health: AFQT scores

3. Measures of education: years of completed schooling
   - Inferences for Union Army data based on whether in school at time of census

Previous research

1-net nutrition (disease and food) affects physical and mental development and health

-Dutch famine (), influenza ()

-improved caloric intake (Fogel & Costa)

2-disease eradication and nutritional interventions lead to increases in schooling and in productivity

-hookworm, malaria, HIV/AIDS

-water-born infections

-vitamin, iron supplementation

Data

- Union Army and Gould Samples

- WWII Enlistment records
  - Height, weight: measured
  - schooling completed

- NHANES 1971-2, 1973-4
  - height measured

- NLSY79
  - AFQT and height self-reported

Union Army: Height & Educ.

- Restrict to ages 11-20 at time of 1850 or 1860 census
- Include all ages when height measured, but include age fixed effects
- If in both censuses, use observation where attending school
- Measures of education are...
  1. if in school at time of census
  2. imputed years of education based on (1), plus data on average school attendance in proximate cohorts.
- Measure (2) is in units (years) that are more comparable (we hope) to other studies. But it is a bit more involved to calculate.

Flow of school attendance by age, 1850-60, IPUMS
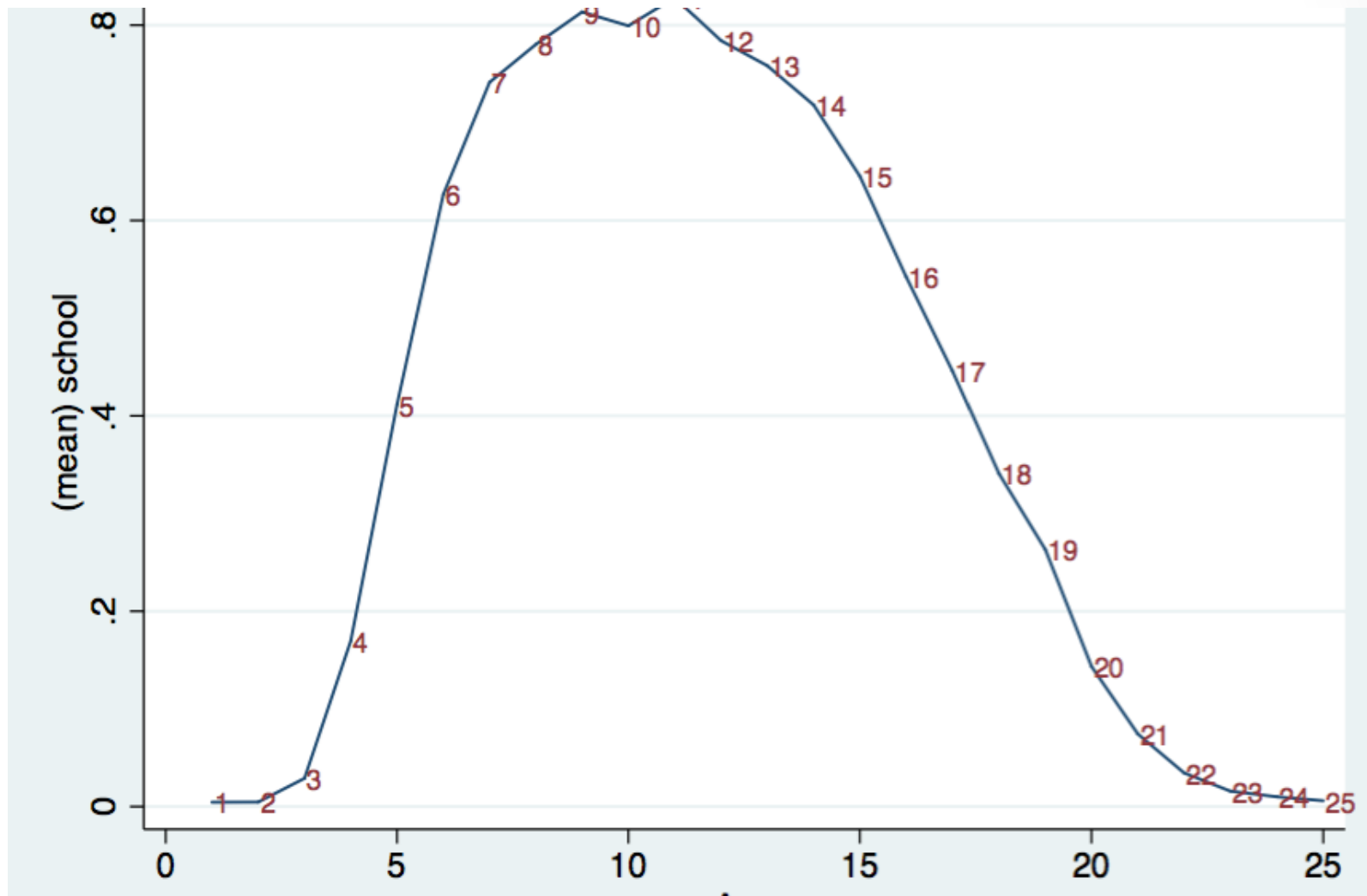
Imputing Attainment

- Cross-section = time series?

- We can form the unconditional expectation of time in school (*e*) for a cohort by simply accumulating the flows (*f*) of school attendance for a cohort. (Margo, 1986, e.g.)

- At age *a*, the expected time is school for the cohort is

$$e_a = \Sigma_{i \leq a} f_i$$

Imputing Attainment

$$e_a = \Sigma_{i \le a} f_i$$

- How much should this expectation of e change if we condition on being observed in school at age a?

- If we had the data, we could calibrate the relationship between eventual attainment and school attendance at a given age. But we don't.

- Instead, model this assuming a "stopping rule".

- A bad assumption when f is still rising with age. ➔ Use a≥11.

Imputing Attainment

$$e_a = \Sigma_{i \le a} f_i$$

- The unconditional expectation of e_a is a weighted average of…
  - those who are "still" in school = (a minus starting age)
  - those who have "dropped out" X_a

$$e_a = f_a\,(a-\underline{a}) + (1-f_a)\,X_a$$

- Solve for X_a.

- Big honking assumption, restated:

    if in school at a, then continuously in school since $\underline{a}$.

Estimates using non-Confederate children

a	f	a-<u>a</u>	X
11	.828057	1	0
12	.784127	2	.203499
13	.758432	3	.3945894
14	.718335	4	.7654896
15	.645366	5	1.431017
16	.542469	6	2.233662
17	.446097	7	2.888962
18	.340817	8	3.545546
19	.263237	9	4.014594
20	.143444	10	4.711823
21	.074341	11	5.106602

Imputing Attainment

- A more generic solution would allow for a distribution of school-start ages and for some relaxation of the "continuously in school" / "stopping rule" assumption.

- Two versions:
  - Pool non-Confederate children to estimate f_a
  - Estimate f_a separately by region.

Union Army Sample, All Ages

Dependent Variable = Education	Dummy=1 if in School	Education 1 (Years)	Education 2 (Years)
Height (cm)	0.266 (0.170)	0.009* (0.004)	0.005* (0.003)
State FE	Y	Y	Y
Age census FE	Y	Y	Y
Age enlistment FE	Y	Y	Y
Log population in town of enlistment	Y	Y	Y
Year census dummy	Y	Y	Y
Standard errors clustered on state	Y	Y	Y

Gould Sample, All Ages, Native-Born

Dependent Variable = height in cm	
No Education (Omitted Variable)	
Limited Common School	-0.734** (0.307)
Good Common School	-0.606** (0.238)
High School	-0.685 (0.492)
City of birth population 50K+ in 1860	Y
Age at measurement FE	Y
State of birth FE	Y
SE clustered on state of birth	Y

UA veterans: Height & Wealth.

- Use subsample linked to 1870 census
- Height (cm) at enlistment (early 1860s)
- Real estate and personal property wealth (1870)
- Standard Census variables as controls: age in 1870, dummies for birth state. Also dummies for year of enlistment.
- LHS variables:
  - Dummy for wealth>0
  - Log_n Total Wealth
  - Log_n Total Wealth, with censored values imputed based on normal approximations truncated at $100.
  - Occupational income score (preliminary)
- Pool "original UA" and preliminary "Urban" sample
- N = c. 3400

UA veterans: Height & Wealth

Dependent Variable = 1870 Wealth	Dummy=1 if Wealth>0	Log Wealth	Log Wealth untruncated
Height (cm)	.003* (.001)	.011* (.004)	.010* (.003)
State of Birth FE	Y	Y	Y
Age 1870 FE	Y	Y	Y
Year of enlistment FE	Y	Y	Y
Sample dummies	Y	Y	Y

UA veterans: Height & Wealth

- Estimates similar in non-urban sample (but depends on whether using the imputation for truncation or not)
- Reweight the UA data to be similar to 1870 IPUMS ages 23-62 and born in Union states by…
  - Place of birth
  - Occupation
  - (Topcode extreme 10% of up-weightings)
- Results not sensitive to re-weighting

- Mixed results for the occupation income score.

- (full table next slide)

UA veterans: Height & Wealth

Sample		Combined			Urban	UA	UA Born in Union				
							State Weight	Division Weight	Region Weight	Occupation Weight	Occ Group Weight
Weight Birthplace Dummy		Region Dummy	No Weight Division Dummy	State Dummy	No Weight State Dummy	No Weight State Dummy	State Dummy	State Dummy	State Dummy	State Dummy	State Dummy
Wealth Dummy	Coefficient	.003	.003	.003	.002	.003	.003	.001	.001	.002	.003
	SE	(.001)	(.001)	(.001)	(.001)	(.001)	(.002)	(.002)	(.002)	(.002)	(.002)
	T-Score	{2.99}	{2.90}	{2.76}	{1.44}	{2.23}	{1.64}	{0.69}	{0.67}	{1.01}	{1.75}
Ln(Wealth)	Coefficient	.011	.010	.010	.007	.012	.013	.008	.008	.008	.012
	SE	(.003)	(.003)	(.003)	(.004)	(.004)	(.006)	(.005)	(.005)	(.004)	(.004)
	T-Score	{3.94}	{3.88}	{3.87}	{1.98}	{3.27}	{2.23}	{1.53}	{1.60}	{2.04}	{2.84}
Ln(Wealth) without generated wealth	Coefficient	.009	.009	.011	.021	.010	.015	.010	.010	.010	.011
	SE	(.004)	(.004)	(.004)	(.011)	(.004)	(.005)	(.005)	(.005)	(.005)	(.005)
	T-Score	{2.50}	{2.66}	{3.06}	{2.01}	{2.65}	{2.87}	{1.94}	{1.91}	{2.22}	{2.16}
Occupation Score (Farmers omitted)	Coefficient	-.004	-.005	-.004	-.062	.042	-.013	.020	.028	-.011	.048
	SE	(.034)	(.034)	(.035)	(.051)	(.051)	(.073)	(.061)	(.061)	(.065)	(.092)
	T-Score	-{0.11}	-{0.15}	-{0.12}	-{1.21}	{0.84}	-{0.17}	{0.33}	{0.46}	-{0.17}	{0.52}
Occupation Score (Farmers included)	Coefficient	-.053	-.051	-.049	-.087	-.032	-.041	-.041	-.039	-.036	-.027
	SE	(.022)	(.022)	(.022)	(.047)	(.025)	(.039)	(.033)	(.033)	(.037)	(.044)
	T-Score	-{2.39}	-{2.32}	-{2.19}	-{1.86}	-{1.27}	-{1.06}	-{1.24}	-{1.18}	-{0.98}	-{0.60}

UA veterans: Height & Wealth

- Compare results for log wealth / cm of height to other studies from more recent samples

- Developed vs developing countries

- Control for education?

20th Century Samples

- White (small samples for non-whites)
- Males: no data on females for older cohorts
- Limit to native born (WWII)
- Ages 20-45 (post-puberty growth spurt):
  - height constant since age 20 (http://www.bmj.com/content/345/bmj.e7331)
  - Drop those older than 45 (not well represented in WWII records)

- Representativeness of the population for these cohorts?
  - Not well-known of WW2 records (compare to 1940 census?)

Measurement Issues

- Height
  - Is height a good marker for health?
  - [40-98 inches]

- Cognition
  - tests comparable over time? Transform into percentiles

- Education
  - In categories:
    - WW2 30% is in the 8 or less: impute 4.5 for time being (look at 1940 census)
  - Absolute or relative?

Mean Years of Education for White Males by Birth Year
Ages 20-45
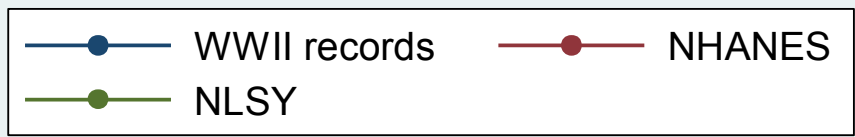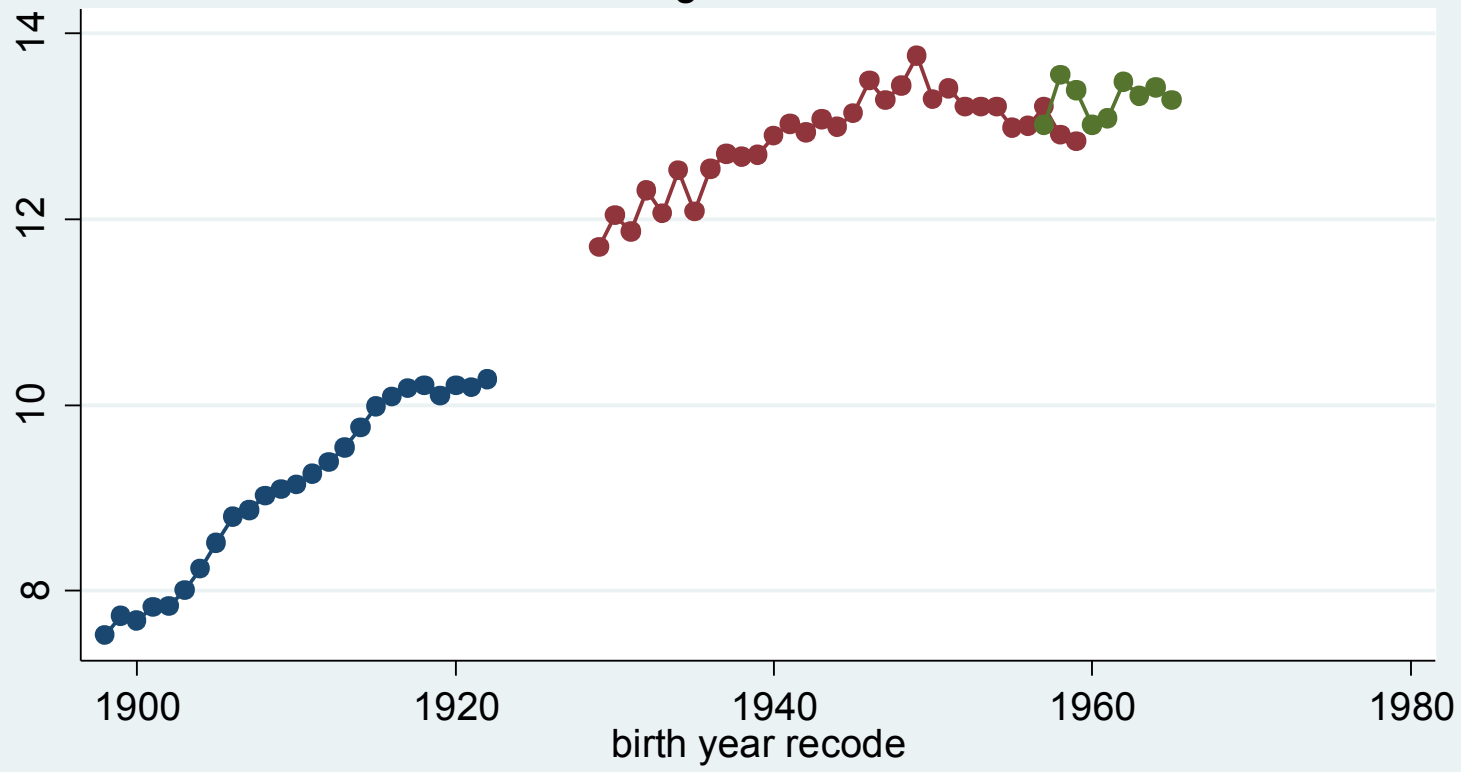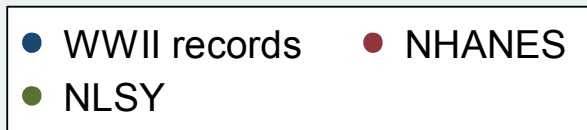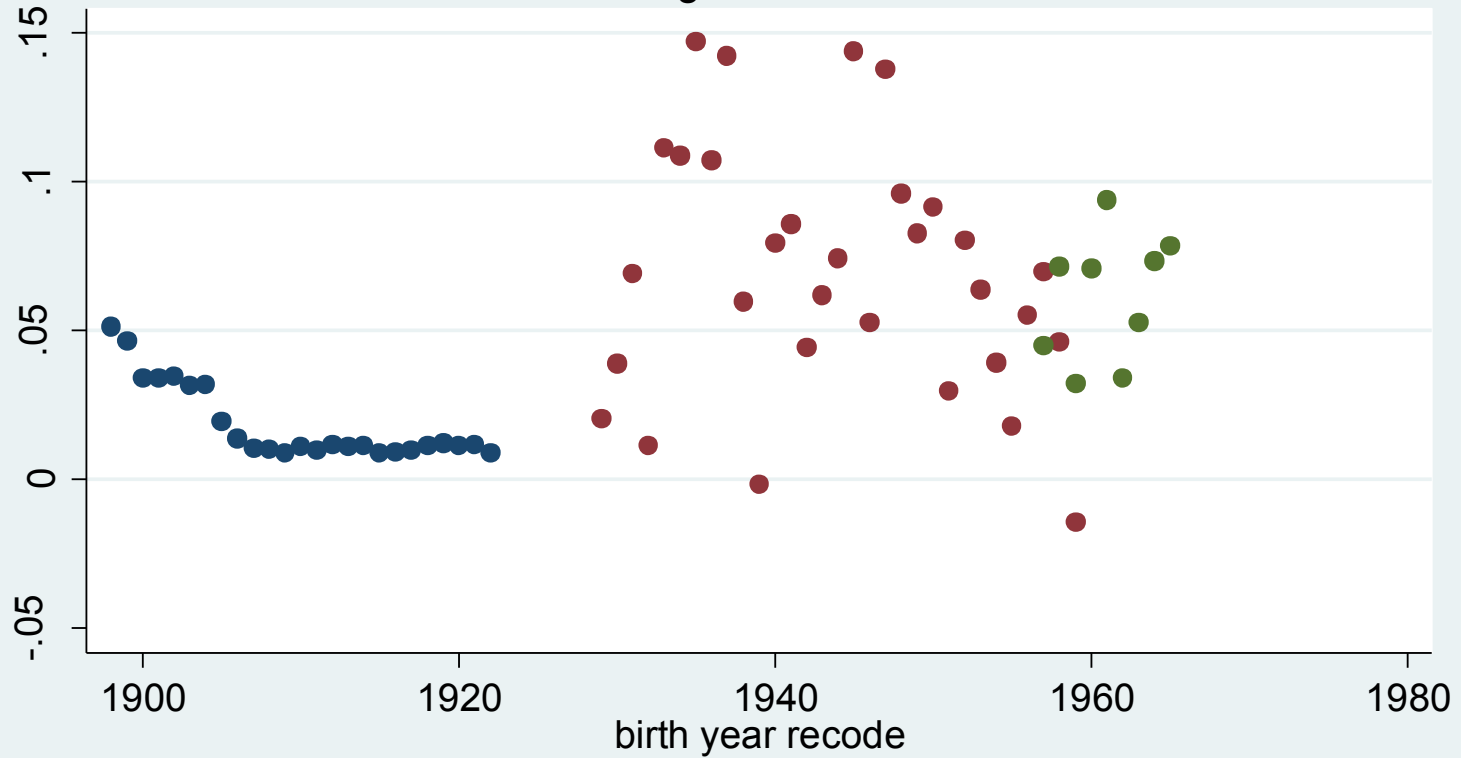
Mean Height for White Males by Birth Year
Ages 20-45
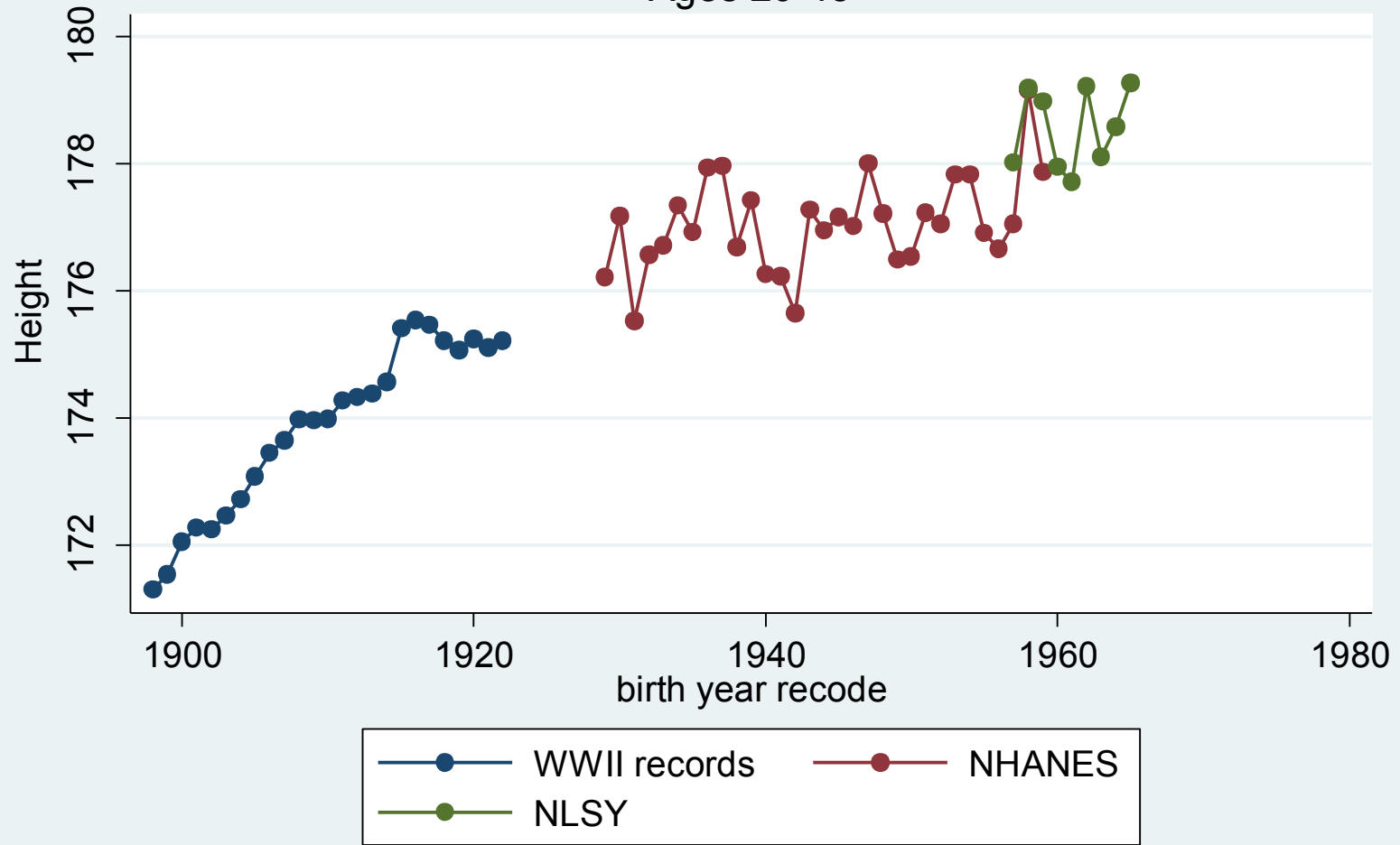
Effect of Height on education for White Males by Birth Year

Ages 20-45

Effect of Height on years of schooling
White native-born males (OLS)

	Basic	cohort dummies	cohort & place of birth dummies
Panel A: WW2 Sample. (N=4,512,722)			
Height (cms)	0.01	0.009	0.01
	[0.000]**	[0.000]**	[0.000]**
R-squared	0	0.03	0.08
Panel B: NHANES 1971-2. N=4,043			
Height (cms)	0.078	0.077	0.074
	[0.006]**	[0.006]**	[0.006]**
R-squared	0.04	0.07	0.14
Panel C: NLSY79. N=2,327			
height (cms)	0.063	0.062	
	[0.007]**	[0.007]**	
R-squared	0.03	0.04	

 * significant at 5%; ** significant at 1%

Determinants of education: Height v. IQ
NLSY79, native born white males

Height (cms)	0.063	0.062	0.007	0.006
	[0.007]**	[0.007]**	[0.005]	[0.005]
AFQT percentile			0.057	0.057
			[0.001]**	[0.001]**
Cohort dummies	no	yes	no	yes
Observations	2327	2327	2216	2216
R-squared	0.03	0.04	0.46	0.46

 * significant at 5%; ** significant at 1%

IQ v height in the WW2 IQ-sample

Height (cms)	0.001	0.001	0.001	0.002	0.002	0.002
	[0.000]**	[0.000]**	[0.000]**	[0.000]**	[0.000]**	[0.000]**
IQ percentile in sample				0.078	0.079	0.078
				[0.000]**	[0.000]**	[0.000]**
Cohort dummies	no	yes	yes	no	yes	yes
State of birth dummies	no	no	yes	no	no	yes
R-squared	0	0.01	0.06	0.4	0.41	0.42

Education and Longevity at Older Ages

- Second half 20th century: Lleras-Muney

- But what about 19th century?

- We present new evidence

  - Union Army datasets

  - Caveat: the same measures of education had before

    - Whether in school at time of 1850/1860 census

    - Imputed years of education

  - All veterans alive and on the pension rolls in 1900

  - Examine years until death using a Cox proportional hazard model

Education and Odds of Death, Union Army Veterans

	Hazard Ratio	Hazard Ratio	Hazard Ratio
In school	0.911 (0.058)		
Education 1 (years)		0.991 (0.012)	
Education 2 (years)			0.984 (0.016)
Age in 1900	Y	Y	Y
Log(city of enlistment population)	Y	Y	Y
1860 census dummy	Y	Y	Y
State of enlistment FE	Y	Y	Y
Age in 1850/60 census FE	Y	Y	Y

Education and Odds of Death from Stroke, Union Army Veterans

	Hazard Ratio	Hazard Ratio	Hazard Ratio
In school	0.652*** (0.056)		
Education 1 (years)		0.924*** (0.012)	
Education 2 (years)			0.888*** (0.018)
Age in 1900	Y	Y	Y
Log(city of enlistment population)	Y	Y	Y
1860 census dummy	Y	Y	Y
State of enlistment FE	Y	Y	Y
Age in 1850/60 census FE	Y	Y	Y

Education and Odds of Death from Ischemic Heart Disease , Union Army Veterans

	Hazard Ratio	Hazard Ratio	Hazard Ratio
In school	1.488 (0.437)		
Education 1 (years)		1.115** (0.060)	
Education 2 (years)			1.164* (0.101)
Age in 1900	Y	Y	Y
Log(city of enlistment population)	Y	Y	Y
1860 census dummy	Y	Y	Y
State of enlistment FE	Y	Y	Y
Age in 1850/60 census FE	Y	Y	Y