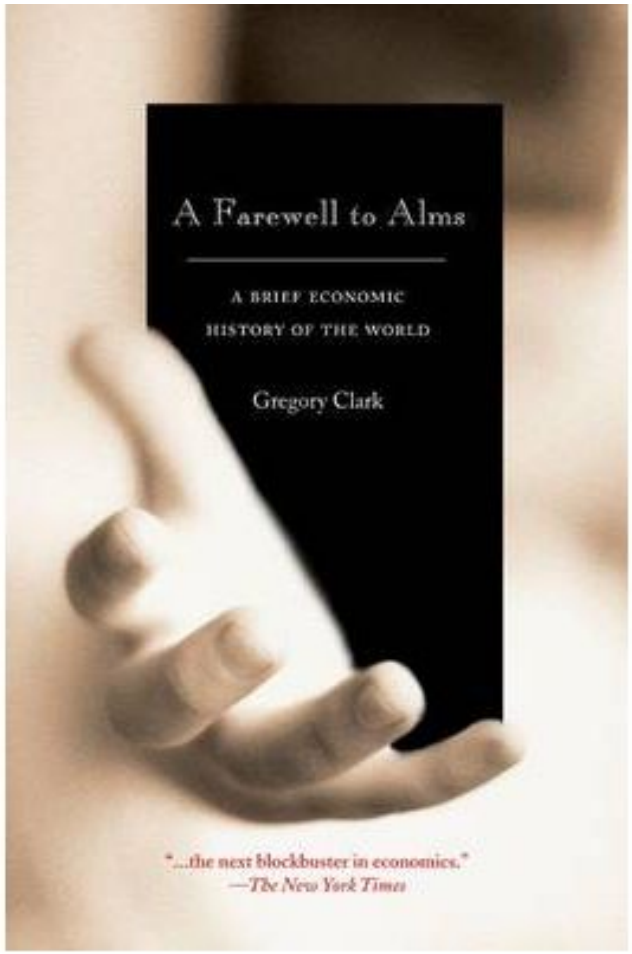
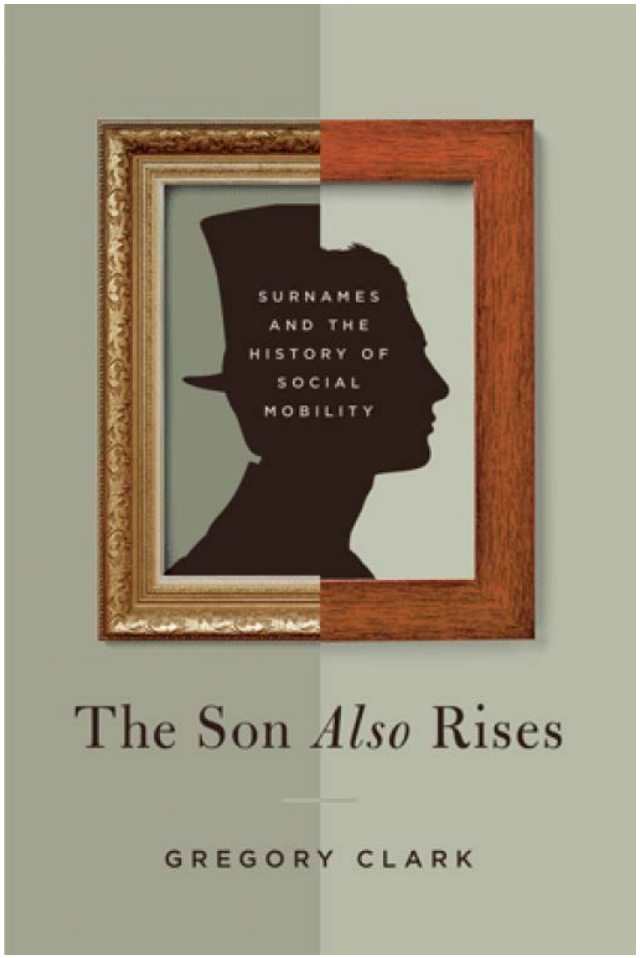
Genetics Determines Social Status: Evidence and Implications from an English Lineage, 1750-2018

Gregory Clark, UC Davis and LSE

Two Books on Economic History – what links them?





A Farewell to Alms (2007)

- Static Malthusian world before 1800 had unexpected dynamic

- "Survival of the Richest"

- Take-over of pre-industrial population, at least in England, by descendants of economically successful 1250-1800

Figure 4: Net Marital Fertility by Wealth Decile, Marriages 1500-1779 and 1780-1879

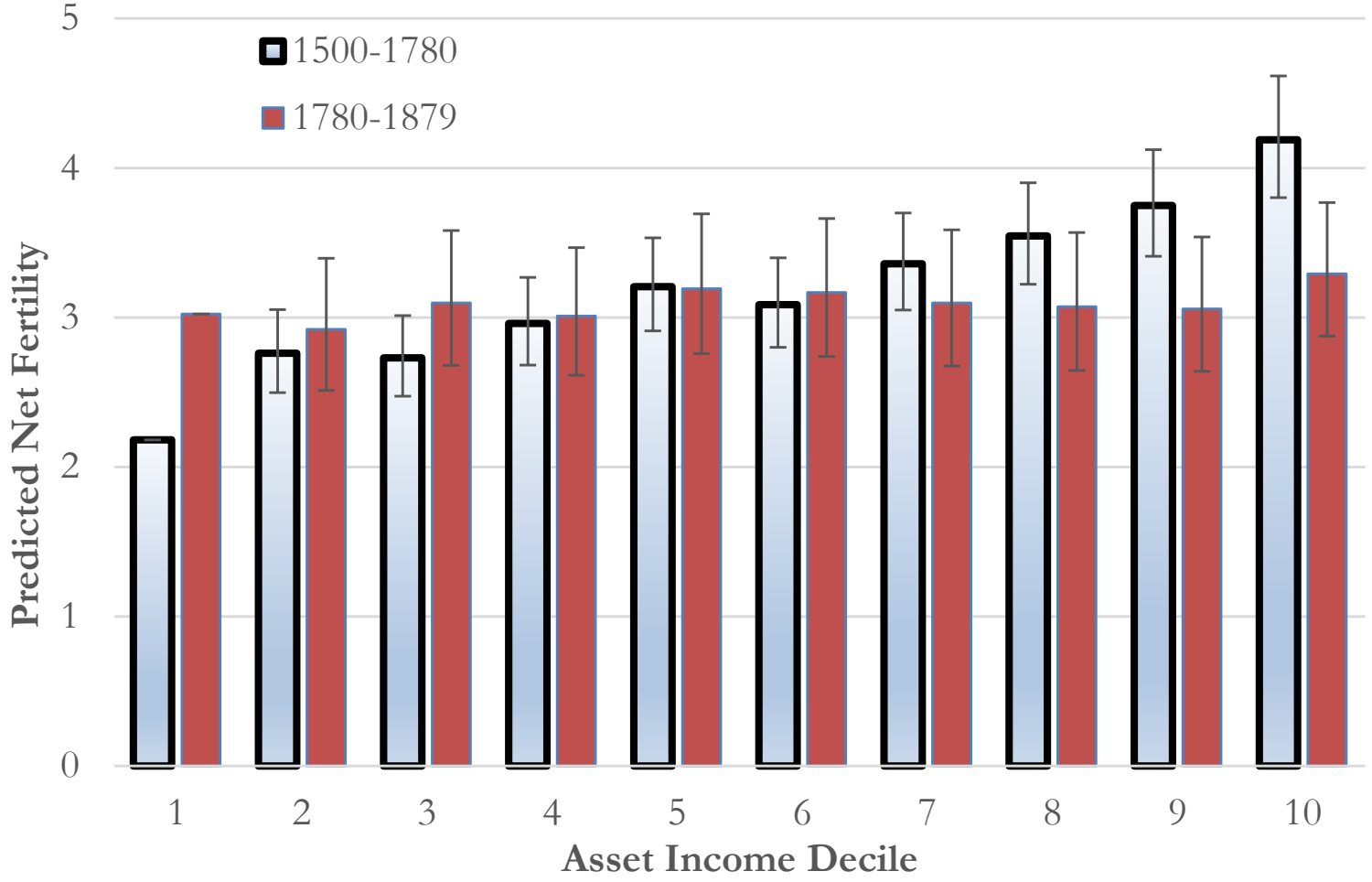
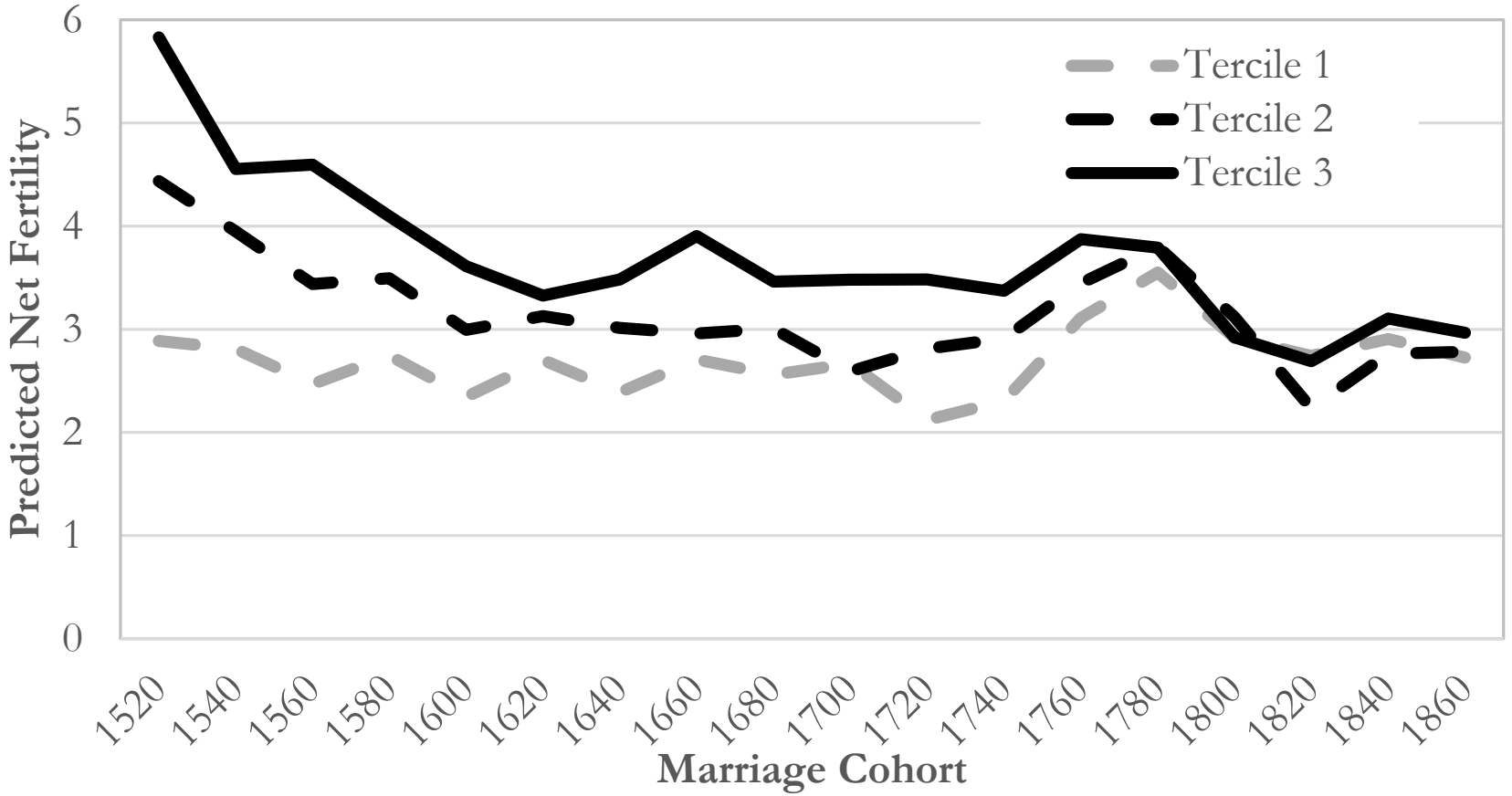


Figure 7: Net Fertility by Terciles, marriage cohorts, 1500-1879



Other Countries also showed this type of pre-industrial demography

- Cameron Campbell and James Lee, *China 1600-1850*

- Sweden, Flanders

England – Associated Social Changes 1200-1800

- Decline in interest rates (10% → 5%)
- Decline in violence
- Increase in educational attainment
  (literacy 1% 1300, 60% 1800)
- (Modest) increase in work hours

Unanswered Questions

- Do demographic regimes have lasting impacts on the economic capabilities of populations?

- Are such impacts mainly cultural or genetic?

The Son Also Rises (2014) examines intergenerational transmission of social status

- Long run social intergenerational correlation of status much higher than short run – 0.7-0.8

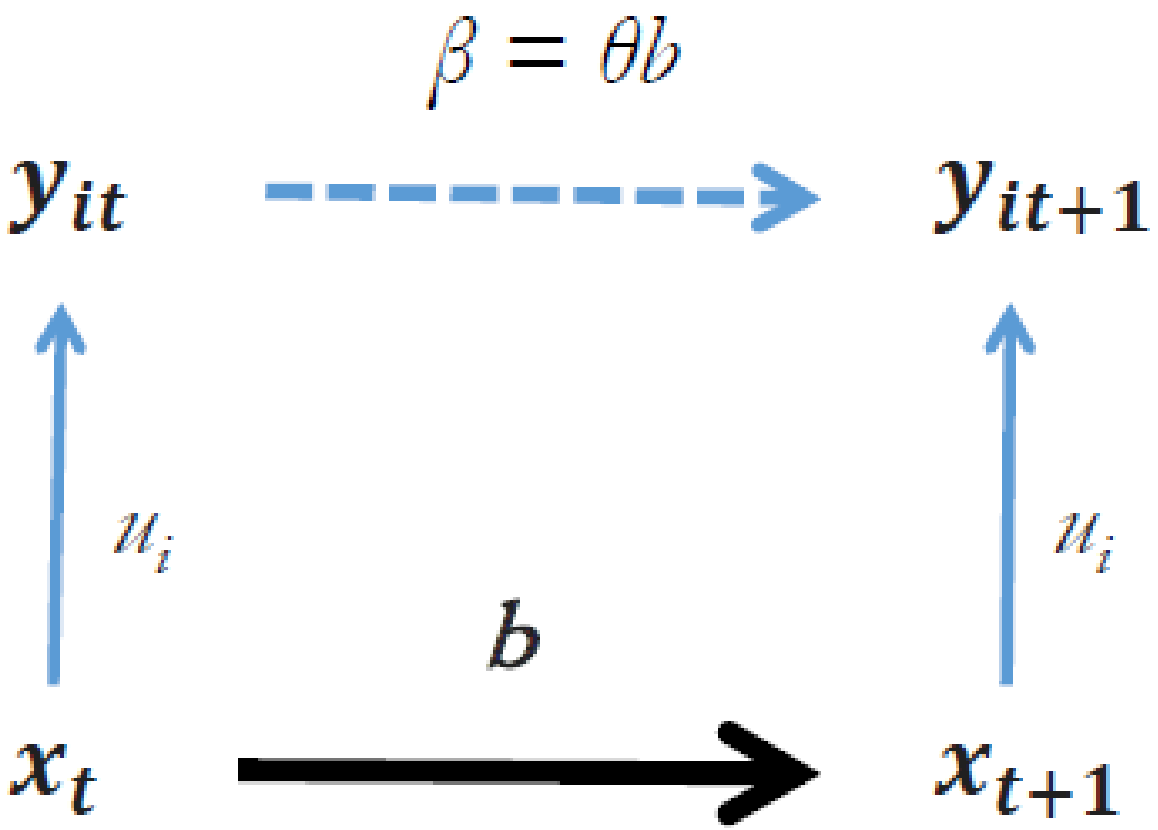
- This rate varies little across societies and time periods – whatever the observed short-run rates

Social Mobility can be well described by a model where

$$y_{it} = x_t + u_{it}$$
$$x_t = bx_{t-1} + e_t$$

- *yi* status phenotypes (observed)

- *x* status genotype (latent)
- b ≈ 0.75


$\beta = \theta b$
y_{it}
y_{it+1}
u_i
u_i
b
x_t
x_{t+1}

Questions

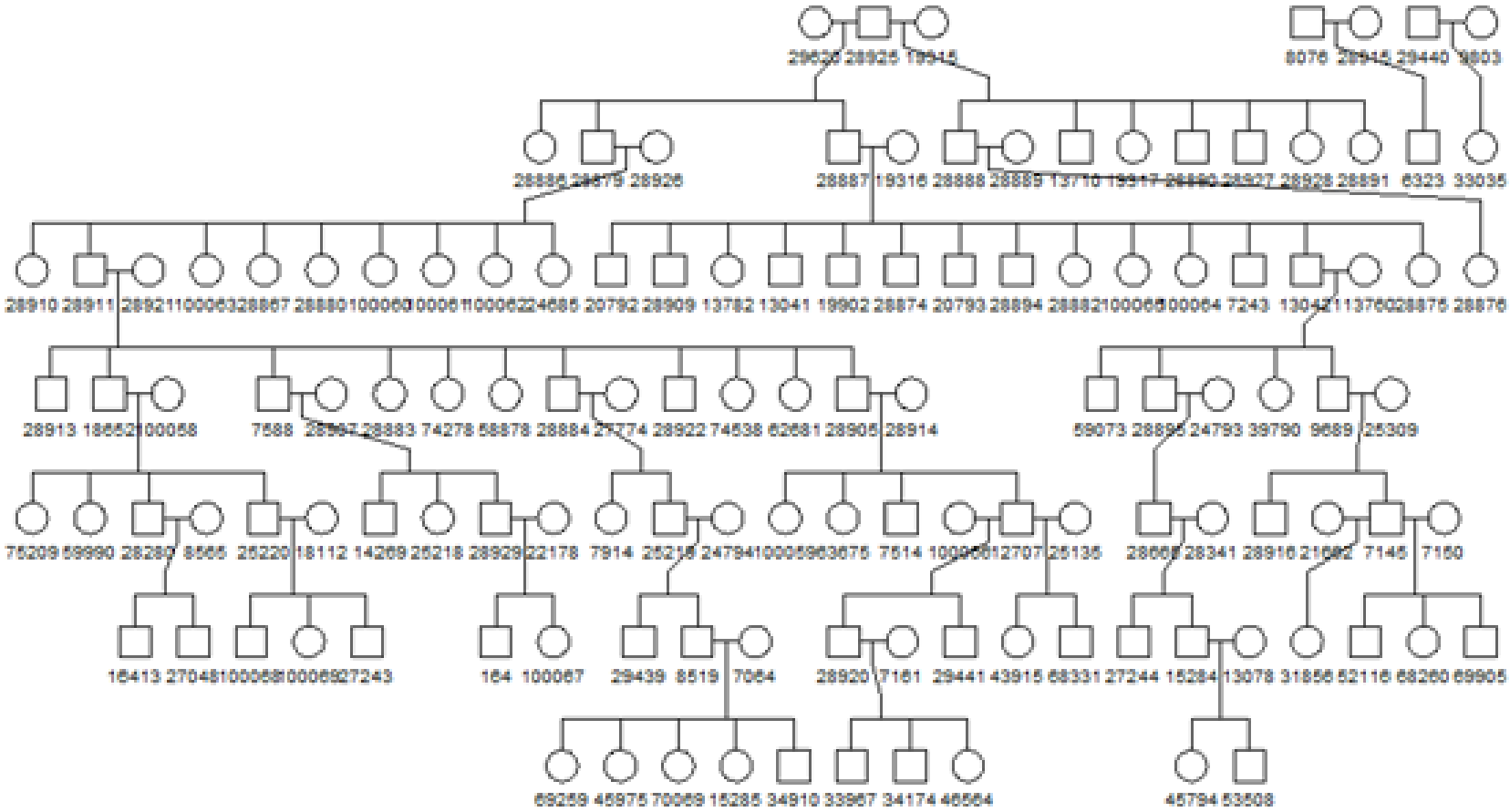
- What is transmitting underlying status so strongly across generations? – How important is genetic transmission?
- Can this transmission be significantly changed by social interventions

Dataset under construction to test this (joint with Neil Cummins, LSE)

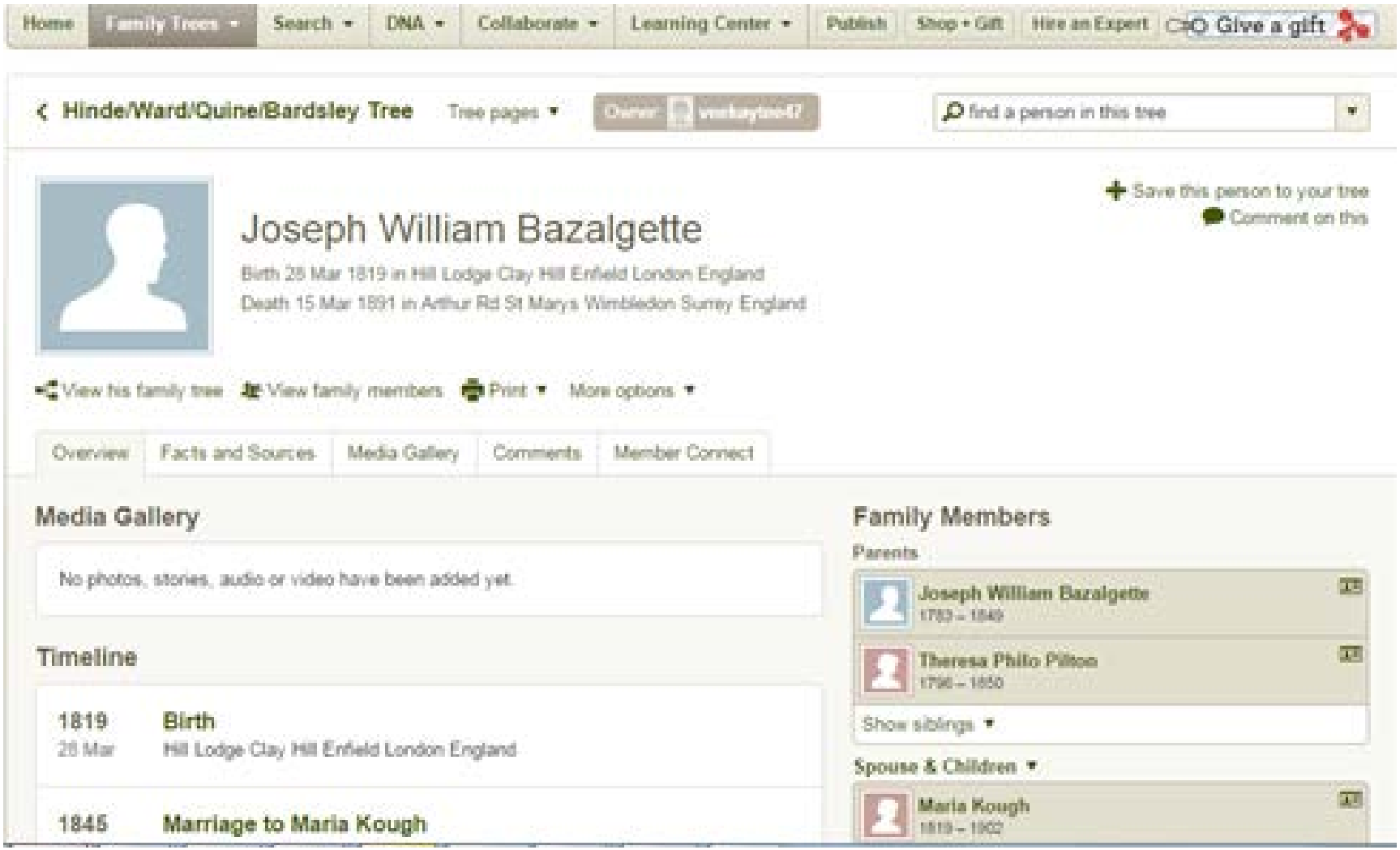
- Lineage of 277,644 people with rare surnames England 1750-2018. We plan to increase size to 350,000-400,000

- Using variety of public data sources we link parents-children across 7 generations

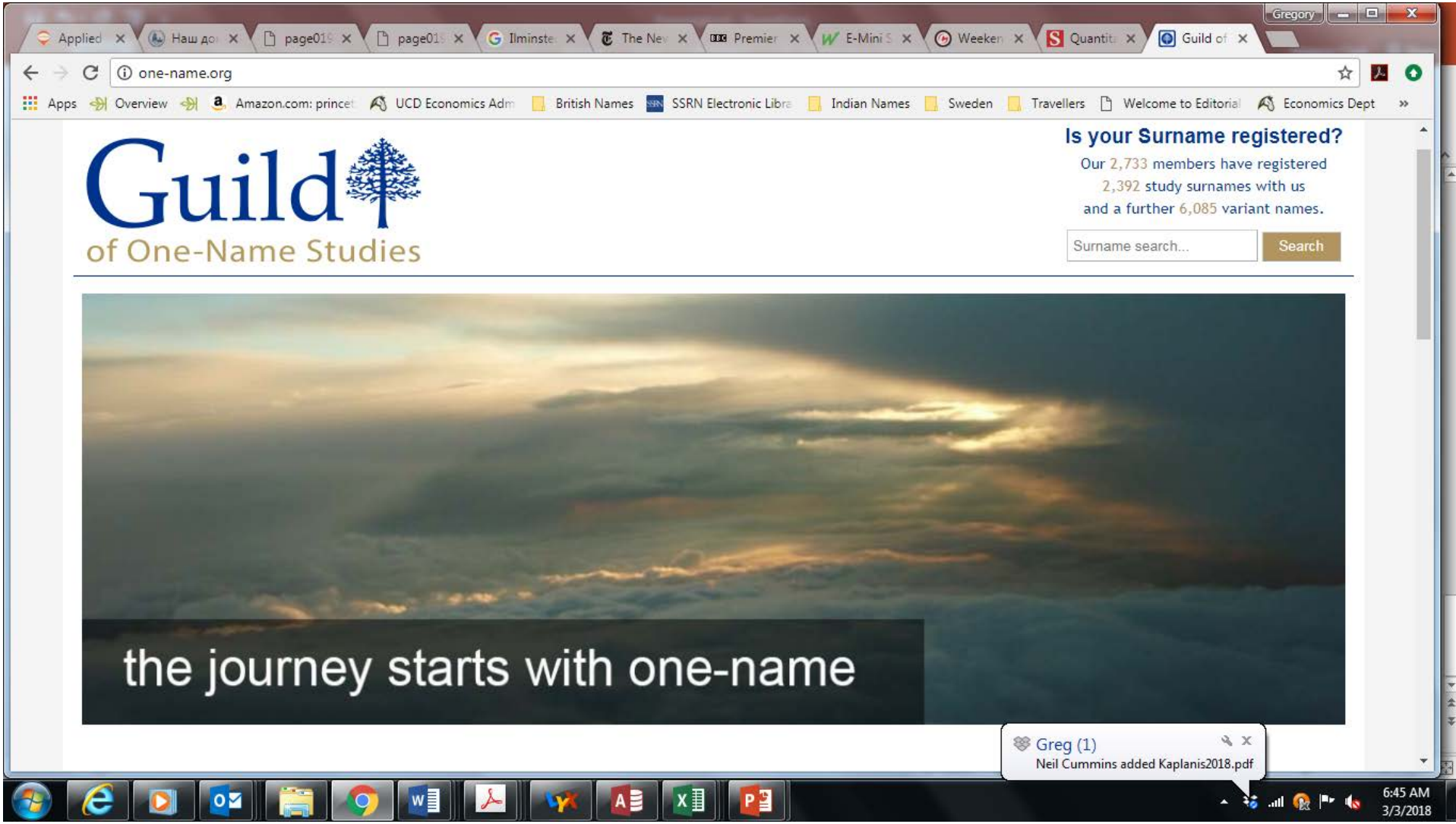
Sample of Database



Crowd Sourcing – Family Trees



Guild of One-Name Studies



Social Outcomes

- **Wealth at Death** – everyone, 1858-2016. Richer families 1799-1857.
- **Adult Occupation** – 1841-1911, 1939
- **Schooling 11-20** – 1851-1911, 1939
- **Education –** Professional Qualifications 1750-1940
- **Life Span** – born 1750-1920
- **Location** – 1841-1911, 1939, 1998-2016
- **House Value –** 2015 (based on address 1999-2012)

Social Outcomes (with sufficient funding £9.25 per copy)

- Birth Certificates 1837-2018 – Parent Occupations
- Marriage Certificates 1837-2018 – Occupations of parties and their fathers
- Death Certificates 1837-2018 – Occupation, cause of death

Data so Far

Variable	Count
All	277,644
Lifespan	111,313
Wealth at Death	42,502
Higher Education (males)	32,266
Occupation (males)	27,215
At Work, 11-20	15,320
Place of Death	106,466
Place of Birth and Death	82,798
Complete Fertility (males)	28,127
House Value, 1999	9,113

Emigration?

- 6% of those born in England and Wales die elsewhere
- We try to follow also migrants
- For those emigrating to Australia and New Zealand, Canada and the USA (75% of emigrants) we can get measures of occupations till 1983 (in Australia)

Is the pattern of social status correlations across relatives consistent with additive genetic inheritance?

- Here in an amazing piece of coincidence we rely on

Fisher, R. A. 1918. "The Correlation between Relatives on the Supposition of Mendelian Inheritance."

Transactions of the Royal Society of Edinburgh, 52: 399-433.

Assumptions of Fisher model seem severe

- The traits in question are controlled by many loci in the genome, each of which makes a small contribution.
- There is an absence of important dominance and epistasis effects.
- Genes and environment are uncorrelated, or the environment has little independent impact on outcomes.

Predicted Correlation of Relatives

Relative to Child	Matching on: Genotype	Phenotype
Parent-Parent	$h^2 m$	r
Mid-parent	h^2	h^2
Single parent	$h^2 \frac{1+m}{2}$	$h^2 \frac{1+r}{2}$
Siblings	$h^2 \frac{1+m}{2}$	$h^2 \frac{1+m}{2}$
Avuncular	$h^2 (\frac{1+m}{2})^2$	$h^2 (\frac{1+m}{2}) \frac{1+r}{2}$
Grandparent	$h^2 (\frac{1+m}{2})^2$	$h^2 (\frac{1+m}{2}) \frac{1+r}{2}$
Cousins	$h^2 (\frac{1+m}{2})^3$	$h^2 (\frac{1+m}{2})^2 \frac{1+r}{2}$
Gt Grandparent	$h^2 (\frac{1+m}{2})^3$	$h^2 (\frac{1+m}{2})^2 \frac{1+r}{2}$
Second Cousins	$h^2 (\frac{1+m}{2})^5$	$h^2 (\frac{1+m}{2})^4 \frac{1+r}{2}$

Long Run Social Mobility

- Depends on $\left(\frac{1+m}{2}\right)$.
- With random matching, correlation long run is 0.5
- For a correlation of 0.8, correlation in genetics of spouses would have to be 0.6

Note

- Additive genetic model with assortative mating on the genotype has the same formal structure across generations as the one derived in ***The Son also Rises***

- In particular social mobility is a first order Markov Process – older generations and collateral relatives play no role

Culture, Resources, Networks

- Sibling correlations should exceed those of parent-child.
- Children grow up in family with same culture, resources, networks
- Not true of parent versus child – given regression to the mean, if this is driven by social environment

Cultural Transmission - Alternative

- $y_{it} = z_t + u_{it}$

z = family culture, shared across siblings

- $z_{it} = bz_{t-1} + e_{it}$

b = underlying long run correlation, e_{it} ensures constant variance in family cultures

Implications

- Parent – child correlation

$$\hat{\beta} = b\frac{\sigma_z^2}{\sigma_z^2 + \sigma_u^2} = \theta b$$

- Sibling Correlation

$$\hat{\rho} = \frac{\sigma_z^2}{\sigma_z^2 + \sigma_u^2} = \theta$$

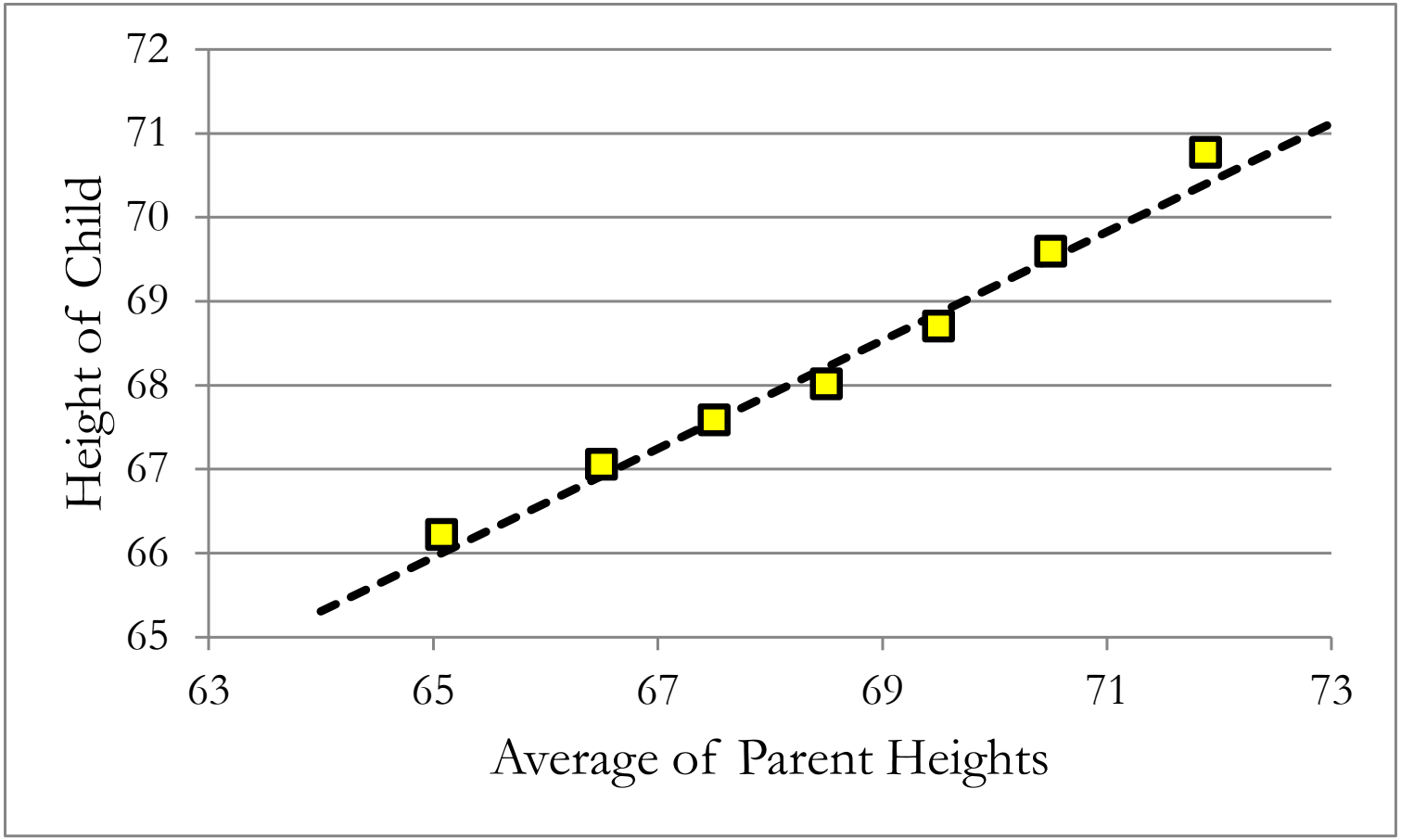
Pattern of Correlations

- Parent-Child θb
- Sibling θ
- Uncle/Aunt θb
- Grandparent θb^2
- Cousins θb^2
- Great Grandparent θb^3

Model of additive genetic transmission – Height Inheritance in modern society

- Process known to be largely genetic
- At least 300 genes known to influence height
- Linearity in regression to the mean

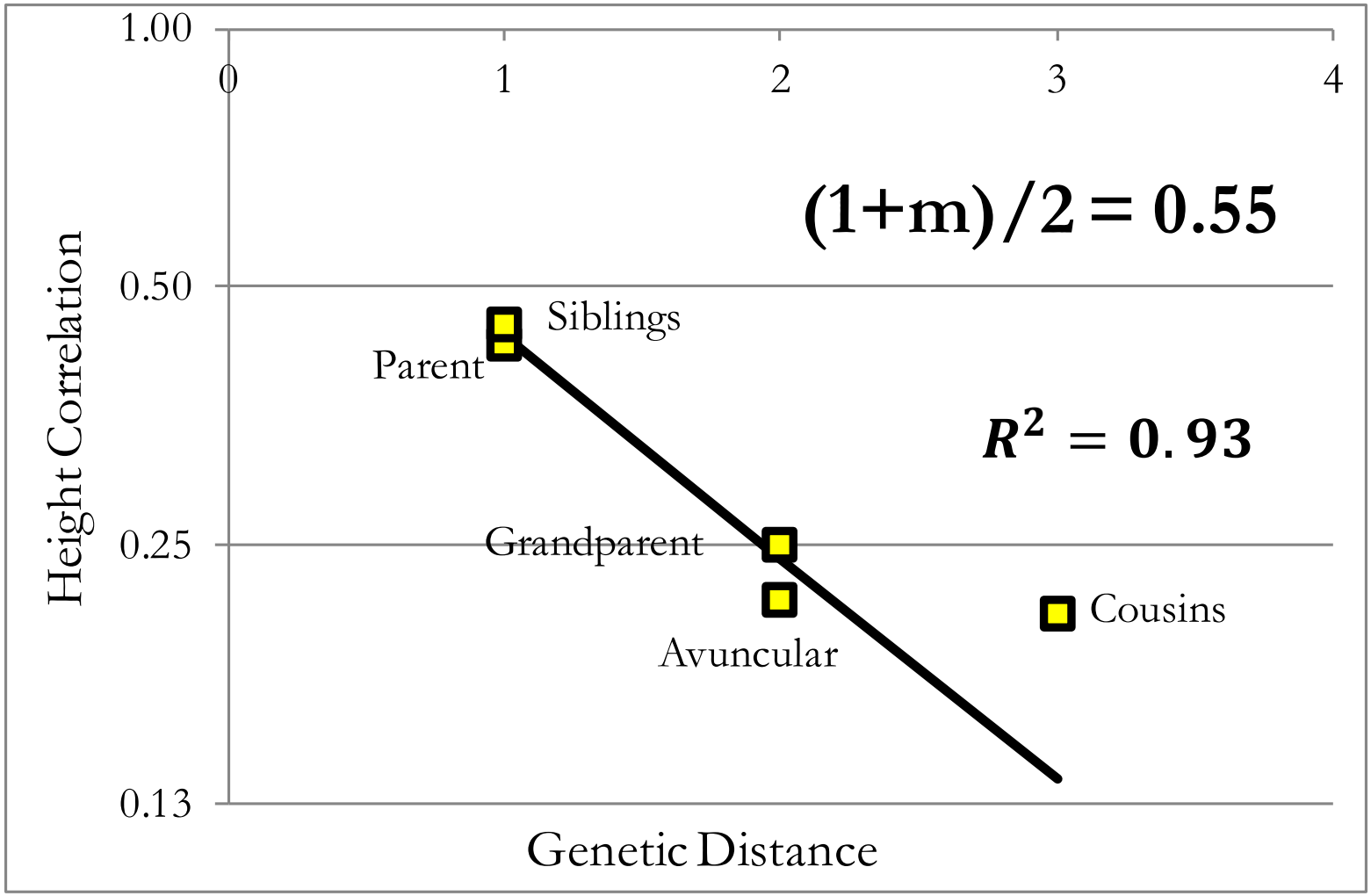
Linearity of Regression to the Mean with Height



Height Correlations Norway, 1984-6

Relation	N	Measured Correlation	Value: Predicted	Fitted
Spouses	24,281	.179	r	.179
Parent-Child	43,613	.430	$h^2 \frac{1+r}{2}$	.430
Siblings	19,168	.453	$h^2 \frac{1+m}{2}$	.412
Grandparent-Child	1,318	.250	$h^2(\frac{1+m}{2})\frac{1+r}{2}$	.243
Avuncular	1,218	.217	$h^2(\frac{1+m}{2})\frac{1+r}{2}$	.243
Cousins	112	.200.	$h^2(\frac{1+m}{2})^2\frac{1+r}{2}$	.100

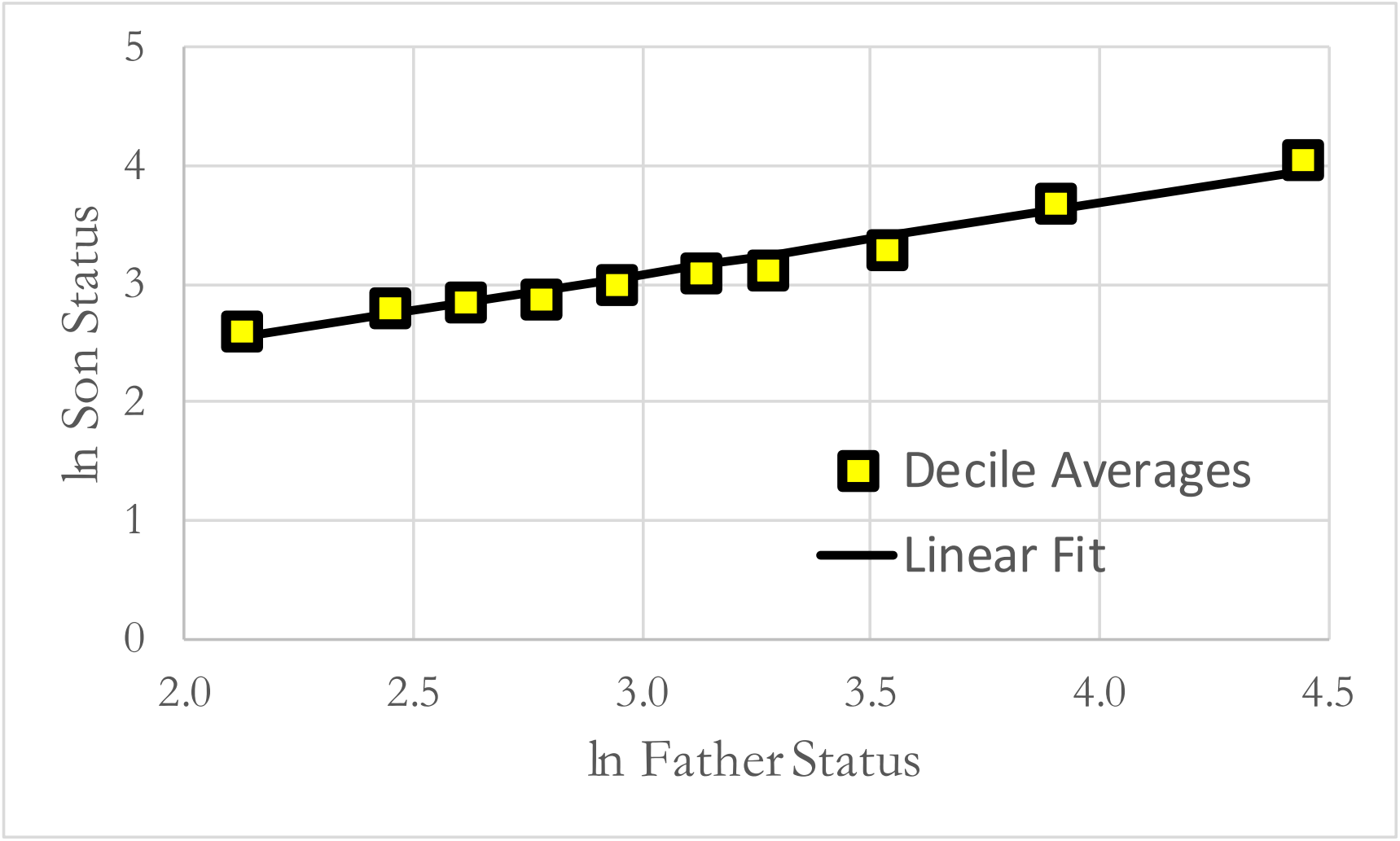
Correlation Pattern Heights



Data on Relatives by Relationship

Relationship	Identical by Descent	All	Higher Ed.	Wealth at Death	Occupation	Age at Death
Father	1/2	89,820	20,611	13,307	16,560	37,662
Brother	1/2	138,651	25,794	22,718	17,878	70,338
Grandfather	1/4	60,897	15,324	9,766	11,870	28,769
Uncle	1/4	171,479	40,396	26,212	30,450	85,324
Gt Grandfather	1/8	44,603	10,096	5,526	7,467	20,610
Gt Uncle	1/8	81,025	16,509	9,376	11,742	37,138
Cousin	1/8	103,665	24,528	16,091	20,498	55,854
Gt Gt Grandfather	1/16	32,173	5,286	2,022	3,700	13,180
Gt Gt Uncle	1/16	61,951	9,227	4,400	6,310	23,959
Gt Gt Gt Grandfather	1/32	22,659	2,516	559	1,834	8,044
2nd Cousin	1/32	98,495	22,544	14,788	19,733	53,753
Gt Gt Gt Uncle	1/32	43,203	4,708	1,620	2,779	13,979
Gt Gt Gt Gt Grandfather	1/64	15,726	1,217	84	1,040	4,948
Gt Gt Gt Gt Uncle	1/64	27,956	2,300	228	1,241	7,670
3rd Cousin	1/128	85,860	15,668	10,558	14,589	44,487
4th Cousin	1/512	71,704	12,272	8,390	12,505	35,002
5th Cousin	1/2048	58,109	10,967	7,887	11,711	31,168

Son Occupational Status relative to Father's Status



Ln Son Wealth relative to Ln Father Wealth, by decile

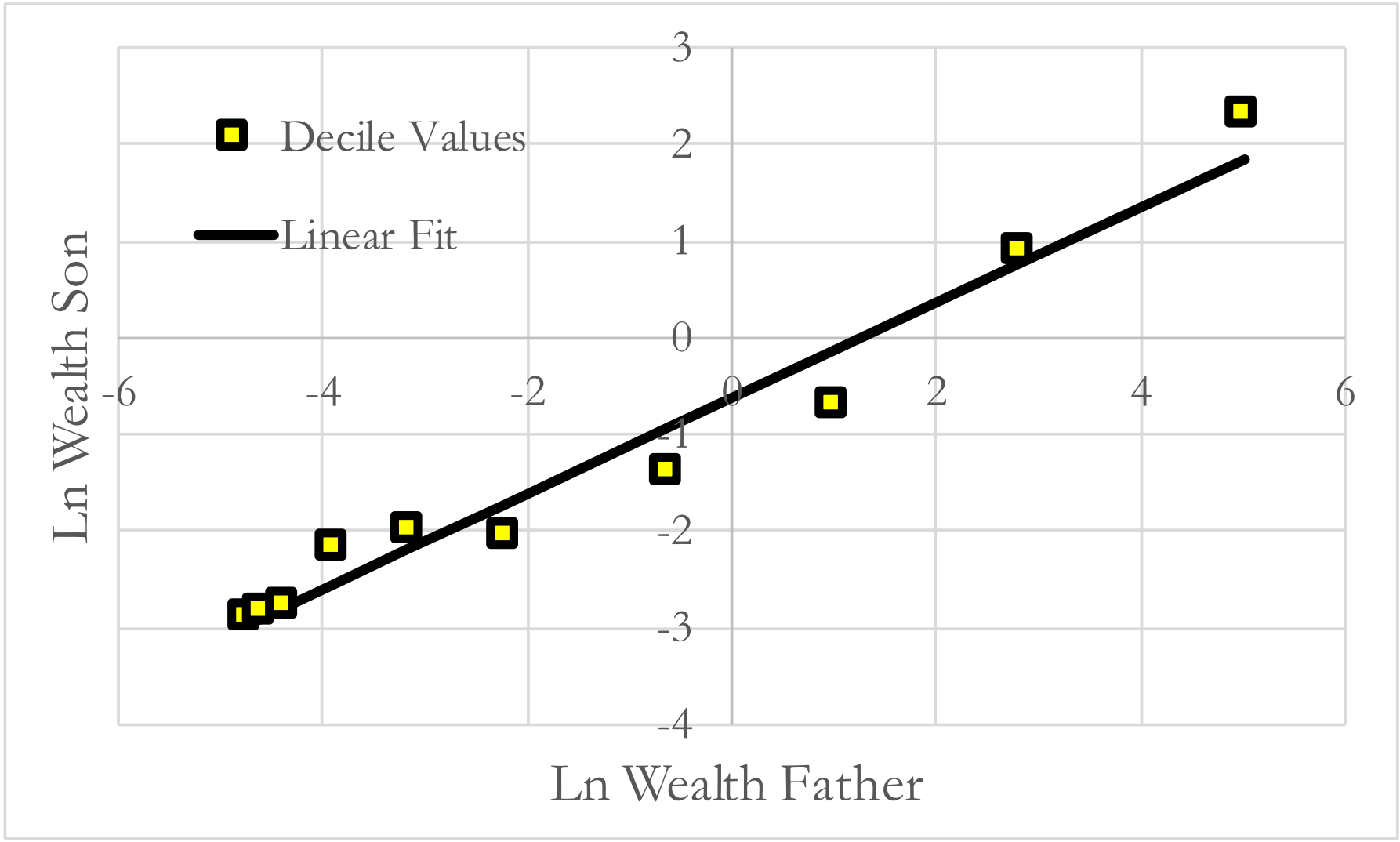
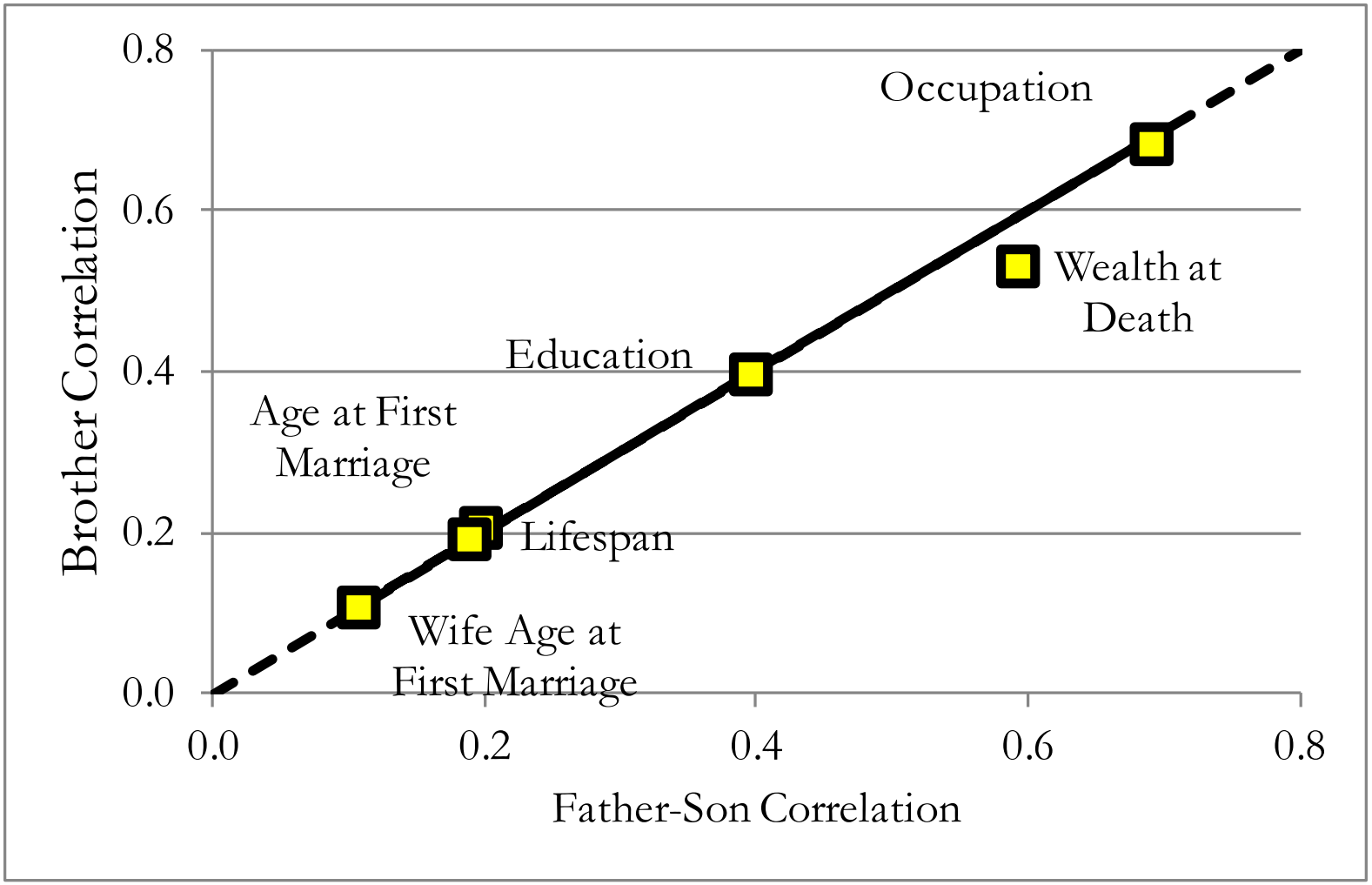


Table 6: Status Correlations

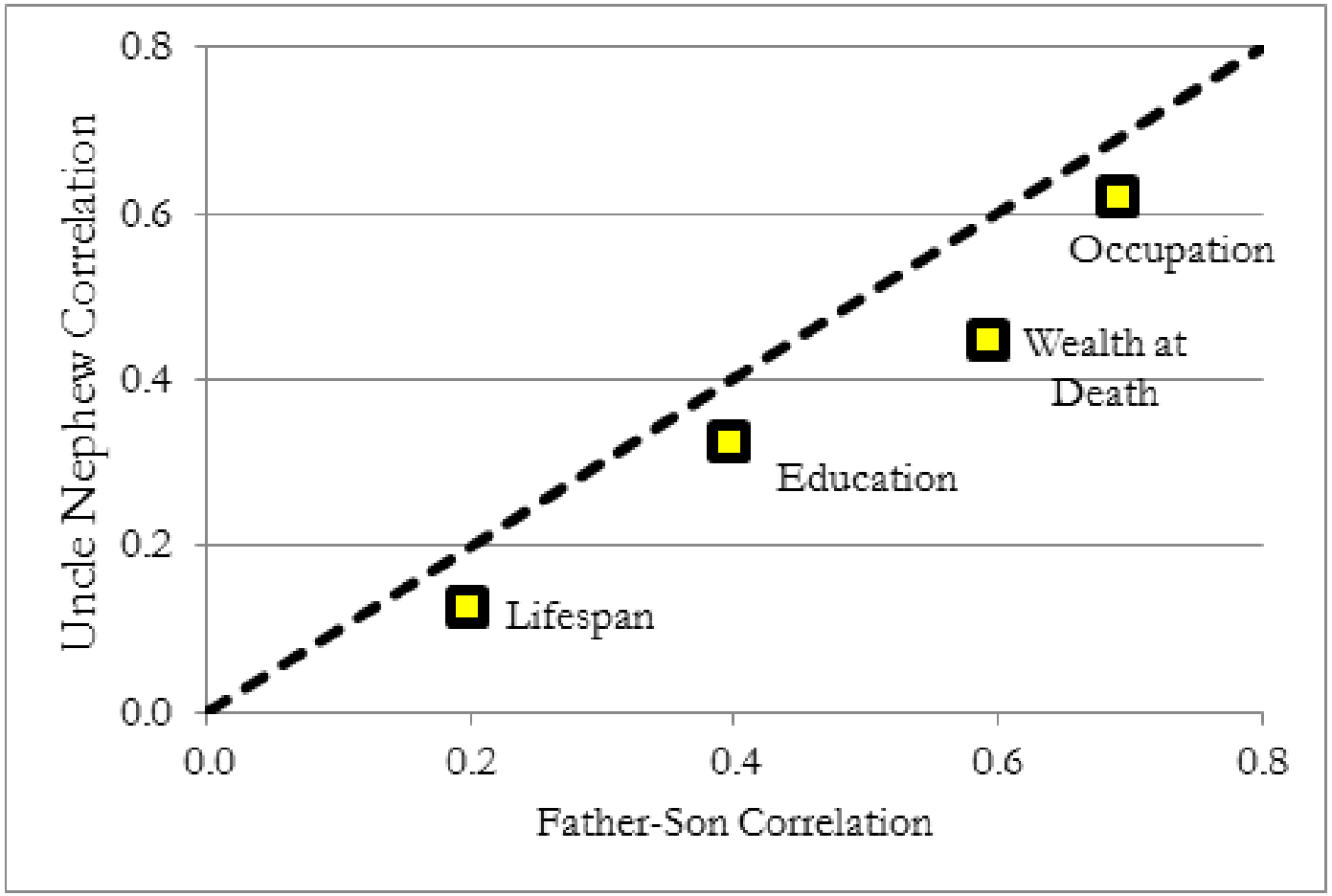
Relationship	Genetic Distance	Wealth	Occupation	Education
Father	1	.592 (.007)	.692 (.006)	.397 (.007)
Brother	1	.531 (.007)	.684 (.006)	.398 (.007)
Grandfather	2	.501 (.009)	.598 (.008)	.305 (.008)
Uncle	2	.447 (.006)	.620 (.005)	.326 (.005)
Gt Grandfather	3	.423 (.012)	.506 (.011)	.219 (.010)
Gt Uncle	3	.402 (.010)	.553 (.008)	.287 (.008)
Cousin	3	.380 (.007)	.605 (.006)	.313 (.007)
Gt Gt Grandfather	4	.391 (.022)	.354 (.016)	.084 (.015)
Gt Gt Uncle	4	.324 (.015)	.449 (.012)	.193 (.011)
Gt Gt Gt Grandfather	5	.379 (.049)	.182 (.025)	.041 (.021)
2nd Cousin	5	.260 (.008)	.446 (.007)	.274 (.007)
Gt Gt Gt Uncle	5	.302 (.030)	.319 (.020)	.087 (.017)
3rd Cousin	7	.136 (.010)	.283 (.009)	.217 (.009)
4th Cousin	9	.096 (.011)	.143 (.009)	.145 (.010)
5th Cousin	11	.030 (.011)	.032 (.009)	.036 (.011)

Note: Pearson Correlations
Standard Errors in Parantheses

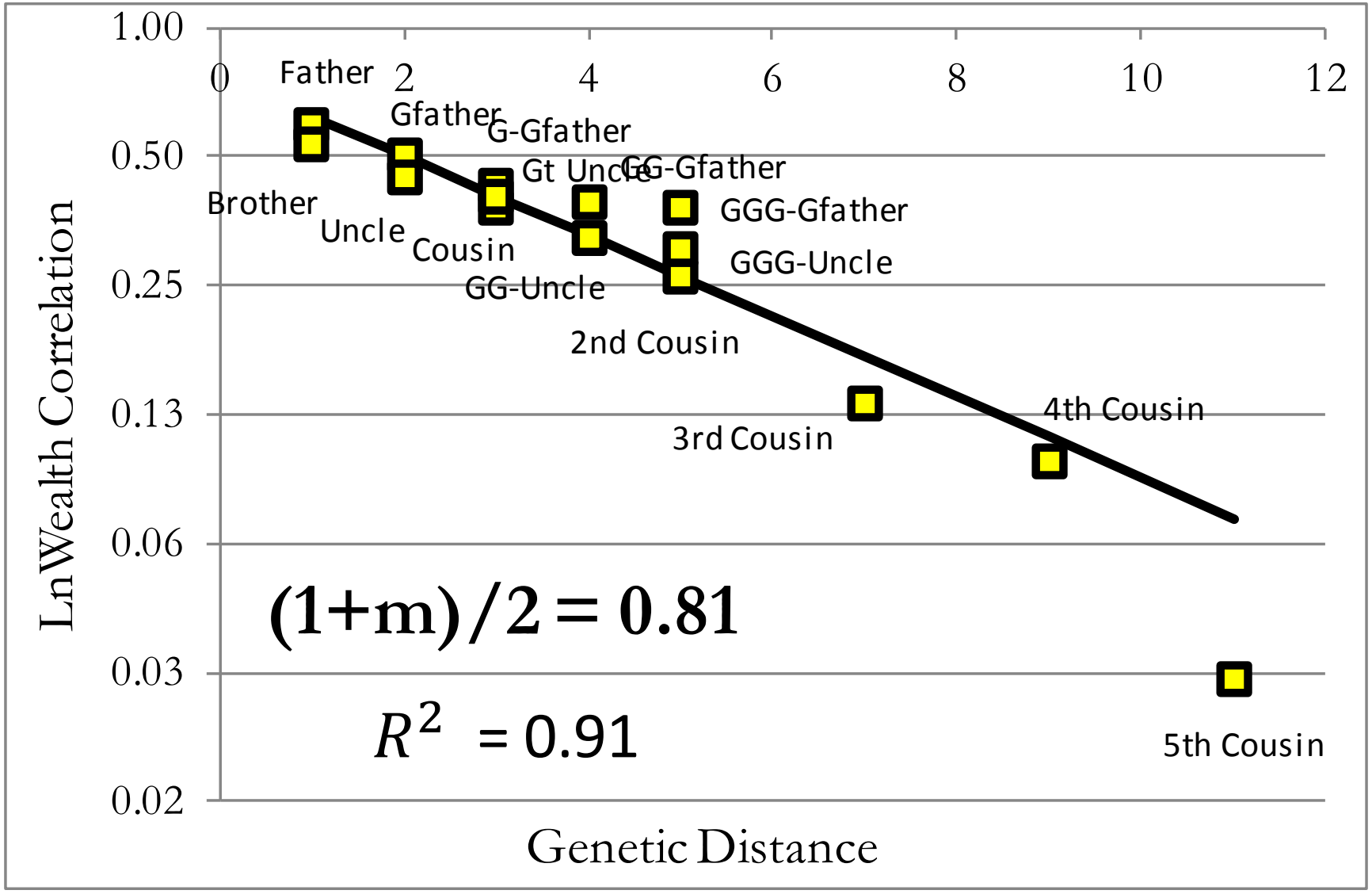
Father-Son vs Brothers



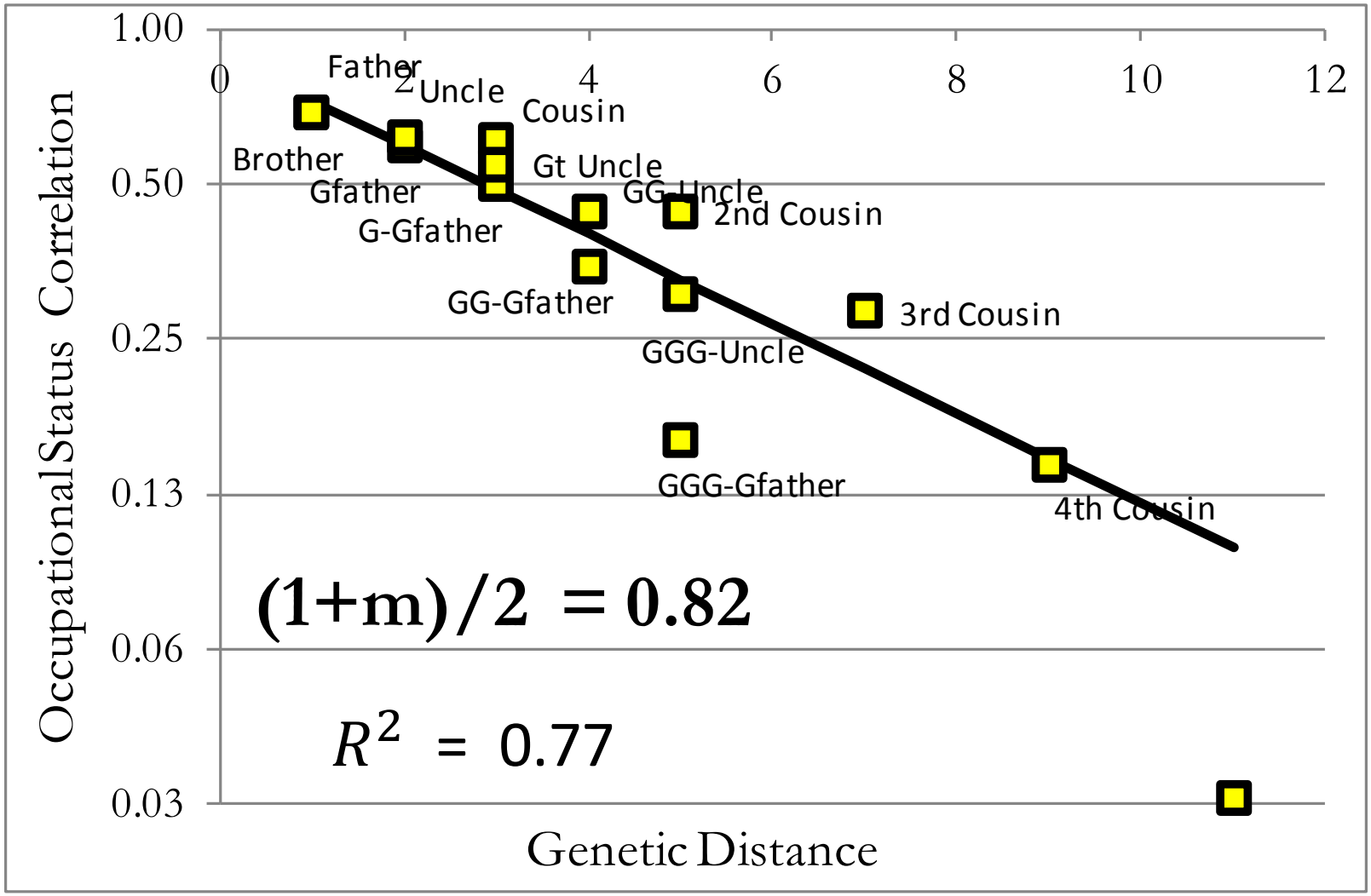
Father-Son versus Uncle-Nephew Correlations



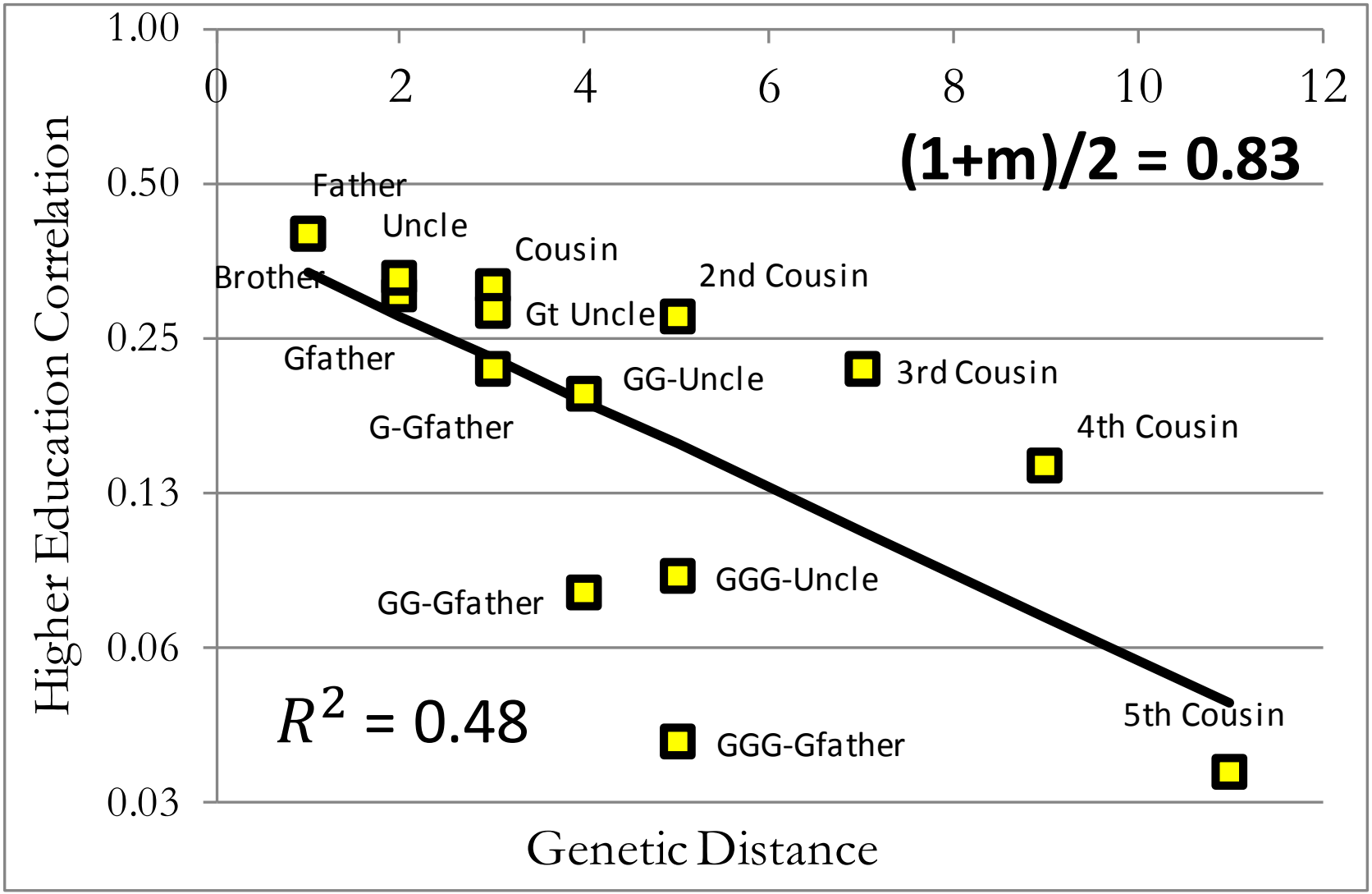
Wealth Correlations and Genetic Distance



Occupational Status and Genetic Distance



Higher Education and Genetic Distance



Is Mating Assortative to the correct degree?

- To get an intergenerational correlation of genotype of 0.8-0.9, then m = 0.6-0.8

- Phenotype assortment is less than this. Years of education 0.4-0.5

Robinson et al. 2017. Genetic Evidence of Assortative Mating in Humans *Nature Human Behaviour*

- Phenotype correlation in years of education 0.4

- Correlation in whole genome predictor of years of education 0.65

Table 9: Instrumental Variable estimates of brother-brother in law correlation

Outcome	OLS Brother-Brother in law	IV lwealth	IV Occupation	IV Education
Ln Wealth	0.413	-	0.905	0.785
	(.021)		(.040)	(.061)
Occupation Rank	0.627	0.927	-	0.838
	(.037)	(.049)		(.055)
Higher education	0.184	0.701	0.603	-
	(.020)	(.032)	(.046)	

If matching is on the genotype then if we estimate the correlation between brothers and brothers in law using IV correlation will increase

$$y_{it} = x_t + u_{it}$$

- *yi* status phenotypes (observed)

- *x* status genotype (latent)
- b ≈ 0.75

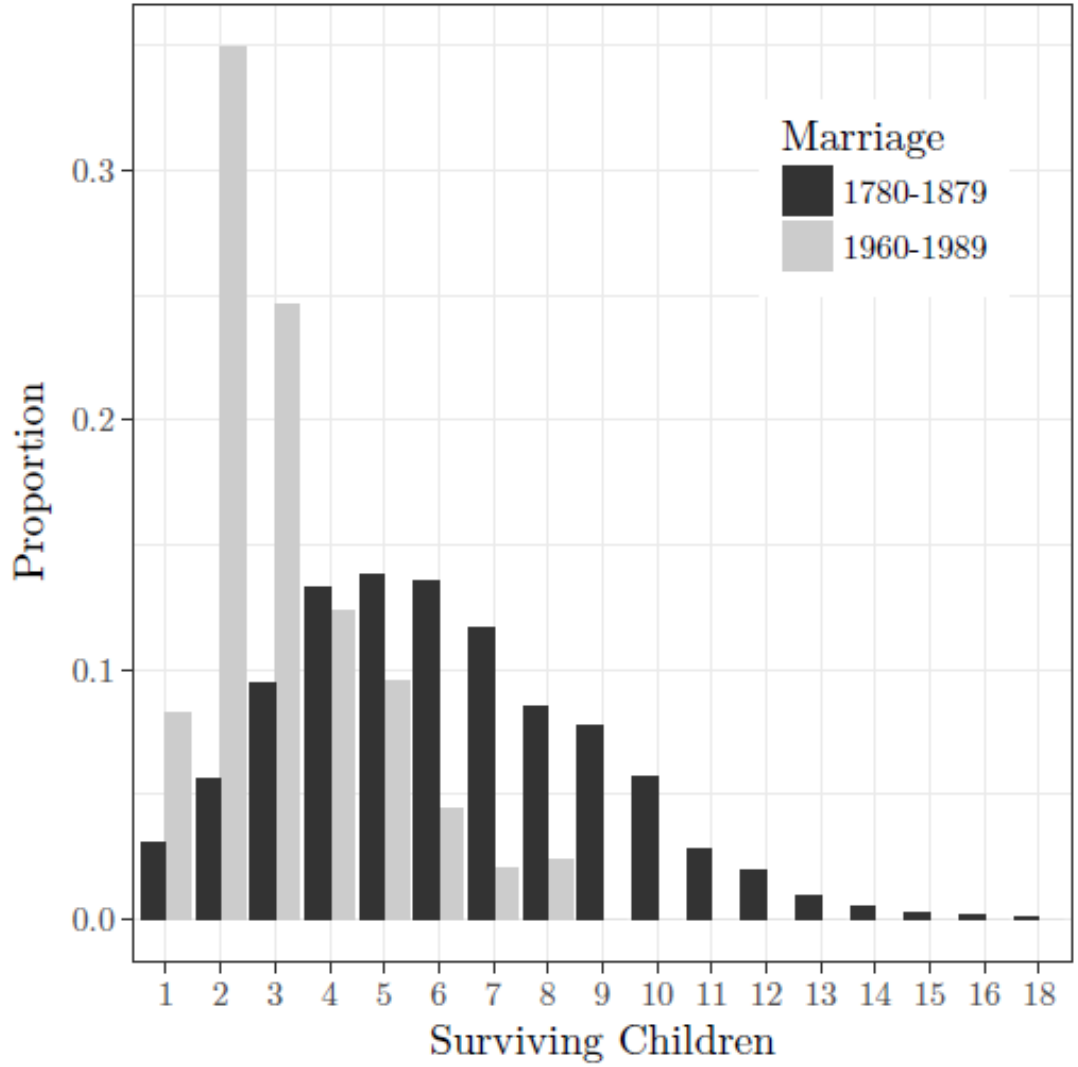
Table 9: Instrumental Variable estimates of brother-brother in law correlation

Outcome	OLS Brother-Brother in law	IV lwealth	IV Occupation	IV Education
Ln Wealth	0.413	-	0.905	0.785
	(.021)		(.040)	(.061)
Occupation Rank	0.627	0.927	-	0.838
	(.037)	(.049)		(.055)
Higher education	0.184	0.701	0.603	-
	(.020)	(.032)	(.046)	

Other Tests

- Family Size, marriages 1780-1879
- Birth Order
- Death of grandparents before birth
- Extent of family social network

Distribution of Completed Family Sizes, England 1780-1879, by child



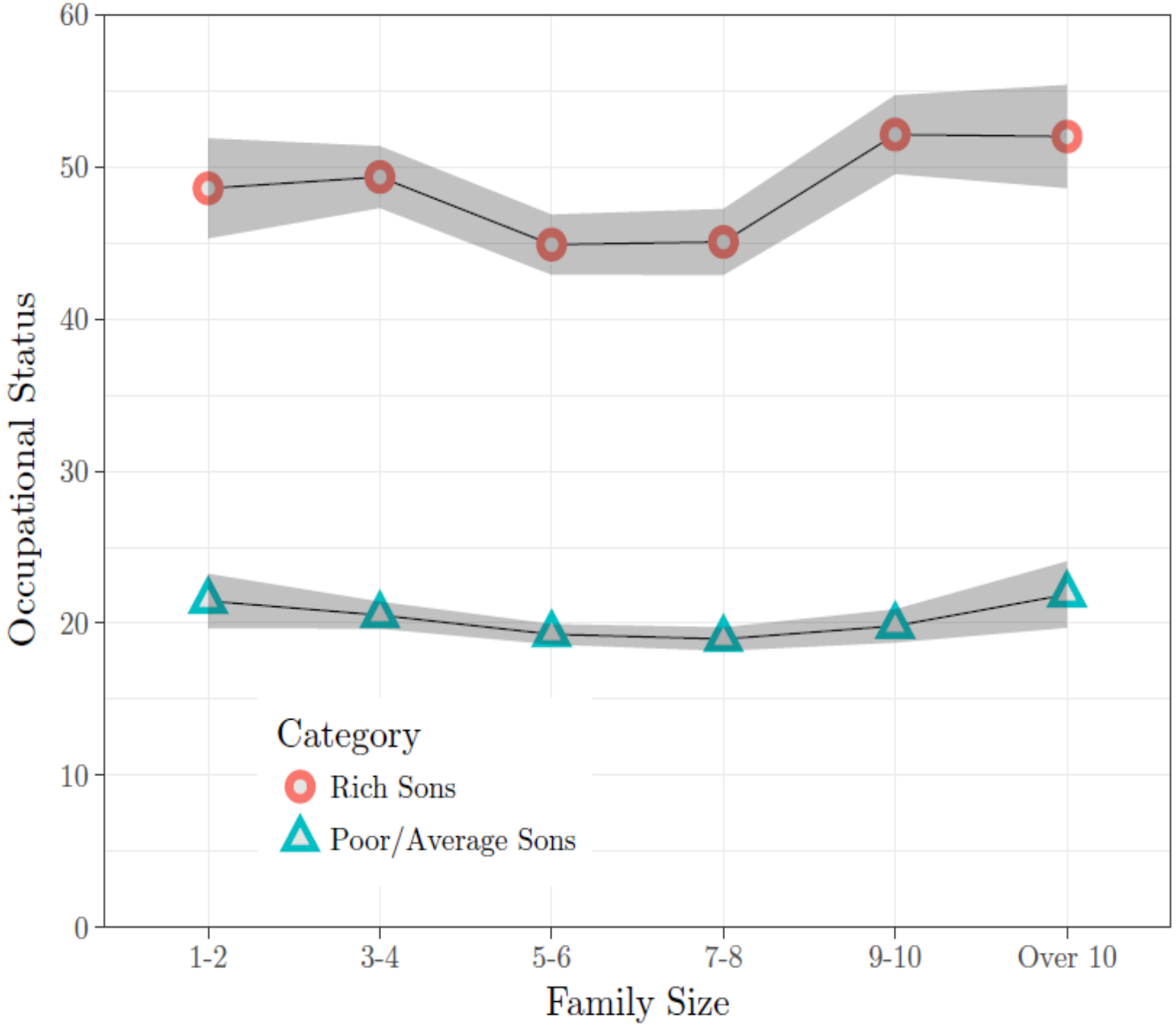


Figure 11: Son Occupational Status, rich and poor lineages, marriages 1780-1879

Notes: 95% Confidence Intervals indicated by shading.

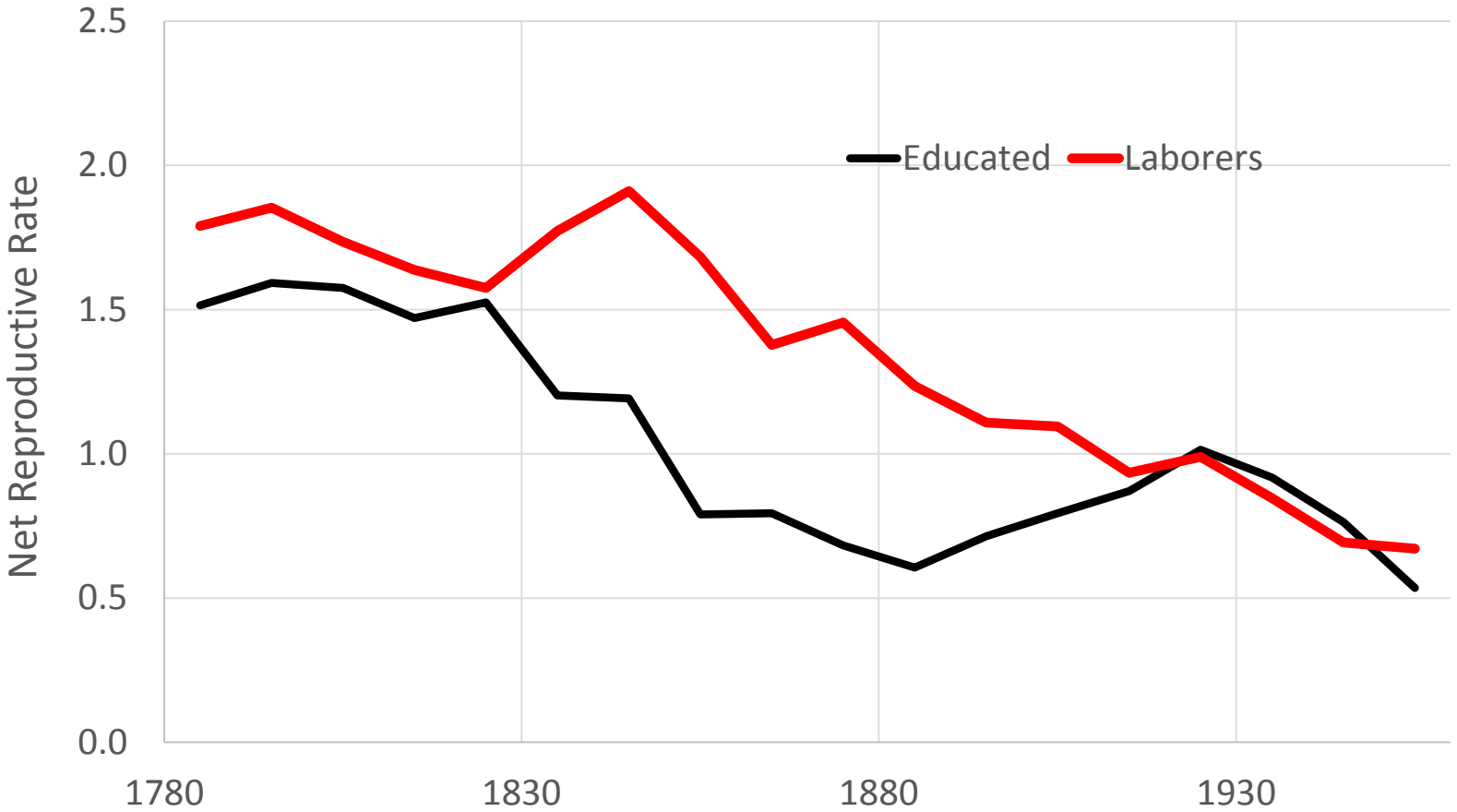
Conclusion

- Differential Reproductive Success of Different Social Classes will change overall economic abilities of the population

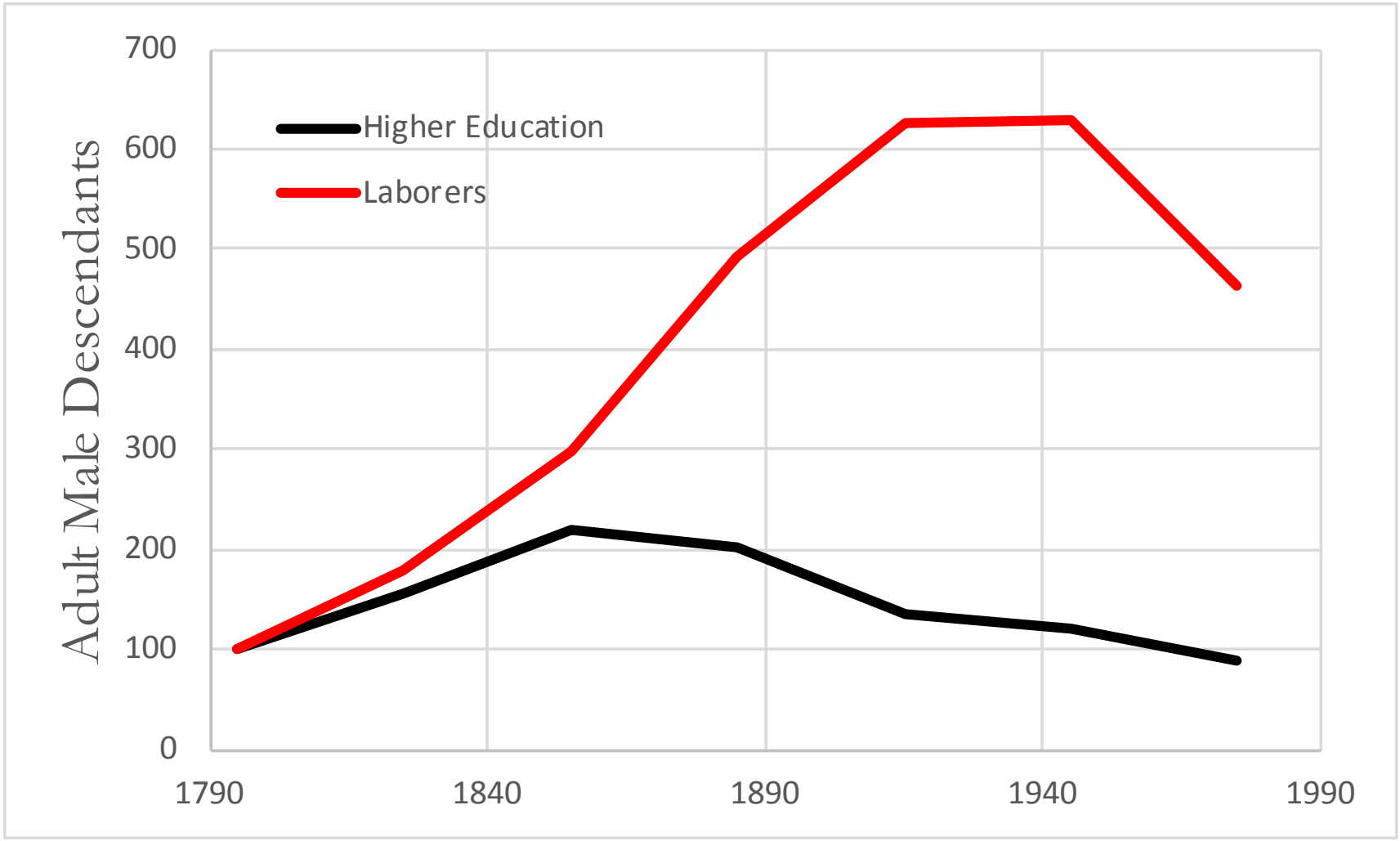
- This would produce a rises in economic abilities in England 1250-1800, in run up to IR

- For men born 1850-1929 there is a strong reversal of the effect, so that average British economic abilities must have declined substantially

Reproductive Success by Birth Decade, Men (higher education) versus Laborers



Cohort Size 1790-1979, Educated Versus Unskilled



How Significant Would Such Effects be?

- Intuition – more effect on distribution of abilities than on the mean

- What matters more for economic outcomes – mean or distribution?

Conclusion

- Interaction of demographic patterns and social classes could have potentially significant effects on the economic capabilities of populations

- England in particular experienced dramatic shift of demographic regimes 1500-2000

Planned Book

- For Whom the Bell Curve Tolls: Culture and Genetics in the History of Human Society