Surnames and a Theory of Social Mobility

Gregory Clark
University of California, Davis

Simple Framework for Measuring Social Mobility

Status Measure y_t : earnings, wealth, years of education, social class, longevity, height (normalized to 0 mean, same variance). Regress

$$y_t = \beta y_{t-1} + v_t$$

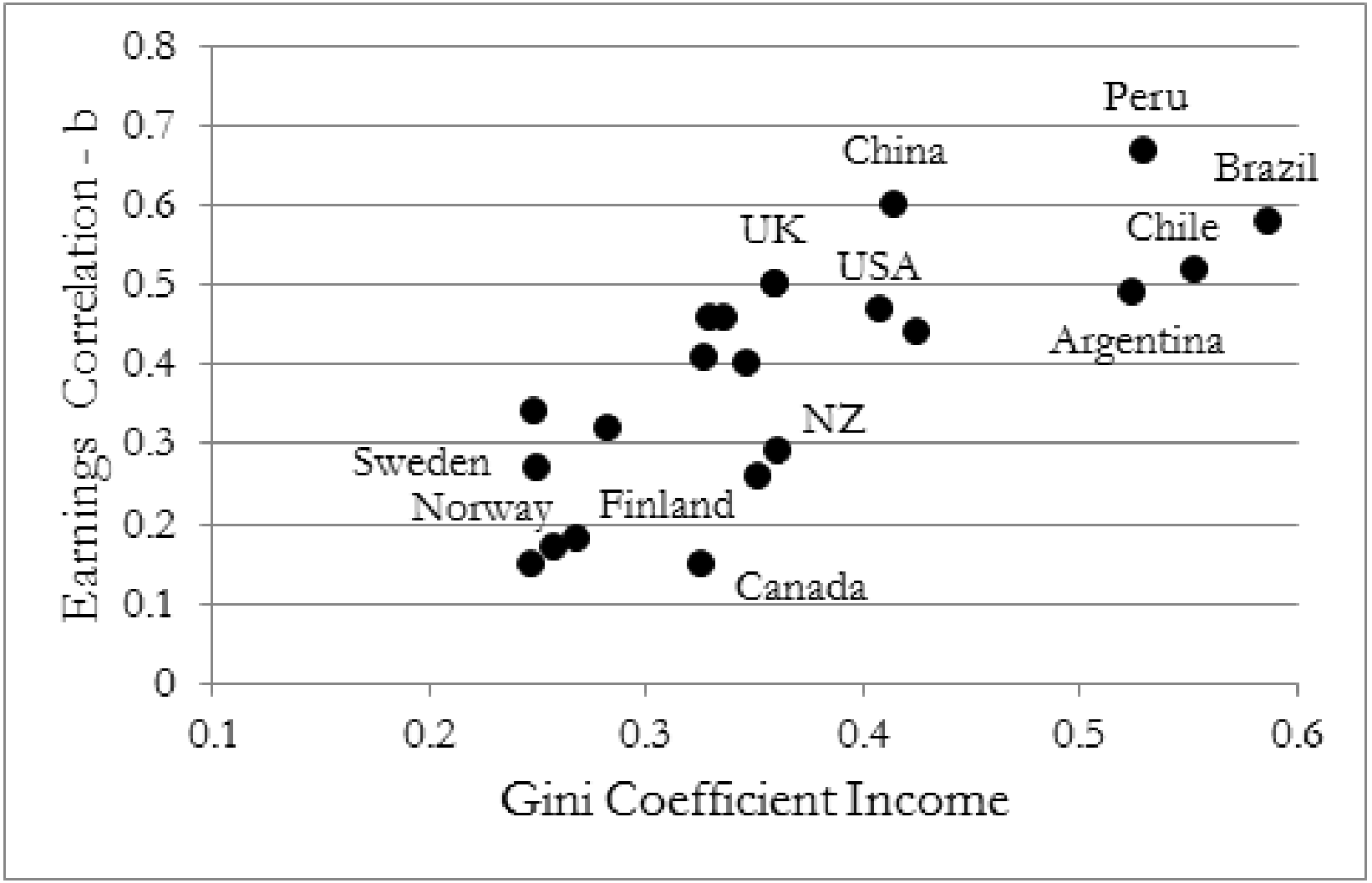
β = intergenerational correlation

t = generation

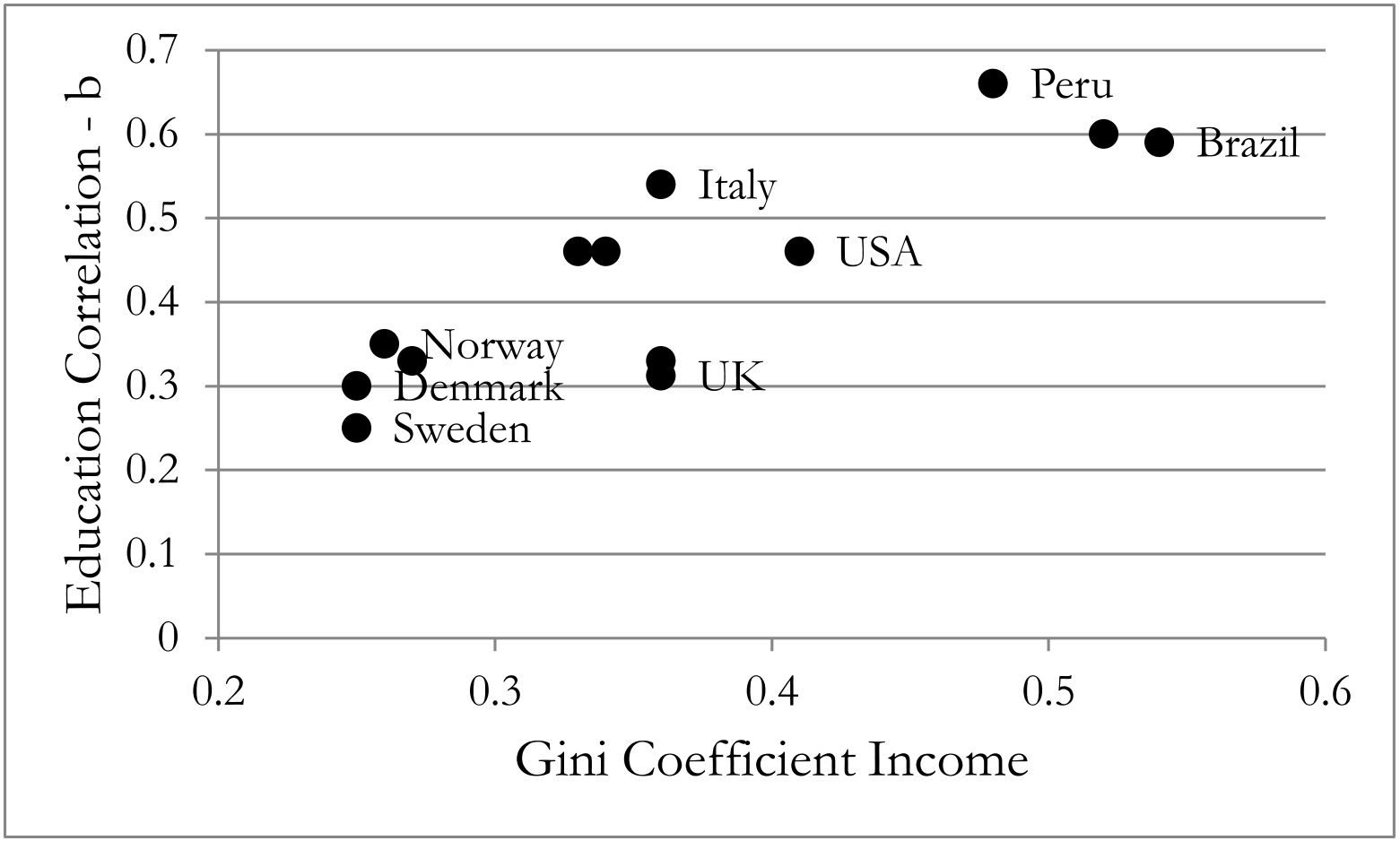
Standard Results with this

- Intergenerational correlations in developed economies of the order of 0.2-0.6 for all measures of status.
- Mobility is rapid. If the process is Markov then within 3-6 generations all earlier status elites, underclasses revert to the mean.
- Mobility rates vary substantially across countries. Inequality is associated with lower mobility rates.

Earnings, β, 2010 (Corak)



Years of Education, β, 1990-2010 (Hertz et al.)



Implications

- Fraction of variance of social status explained by inheritance low – 4% Scandinavia, 22% USA
- Mobility rates must be "too low" in some societies

Recent studies of multiple generations consistently suggest, however, that the process is not Markov. If we estimate

$$y_{t+1} = \beta_1 y_t + \beta_2 y_{t-1} + \beta_3 y_{t-2} + u_{t+1}$$

then $\beta_2 > 0$, $\beta_3 > 0$ and so on.

Even controlling for parents, the status of grandparents, and even great-grandparents is predictive of this generation's status

Surname Method

- Measure social mobility by tracing wealth, income, education averages by surname lineages – e.g. Clark, Smith

- Surnames link us to previous generations though the patriline – in England we can link some people alive now to their ancestors in 1066

- With the high rates of social mobility typically found, common surnames should rapidly lose status information

Measures of Status

- Direct measures of wealth, income, longevity
- Fraction of people bearing a surname who are in high status occupations, or are wealthy – doctor, attorney, member of Parliament, professor, author, probated
- Fraction of people bearing surname who are educated at universities – Oxford, Cambridge, 1200-2012, Uppsala, Lund, 1700-1954

Intergeneration Correlation from Surnames

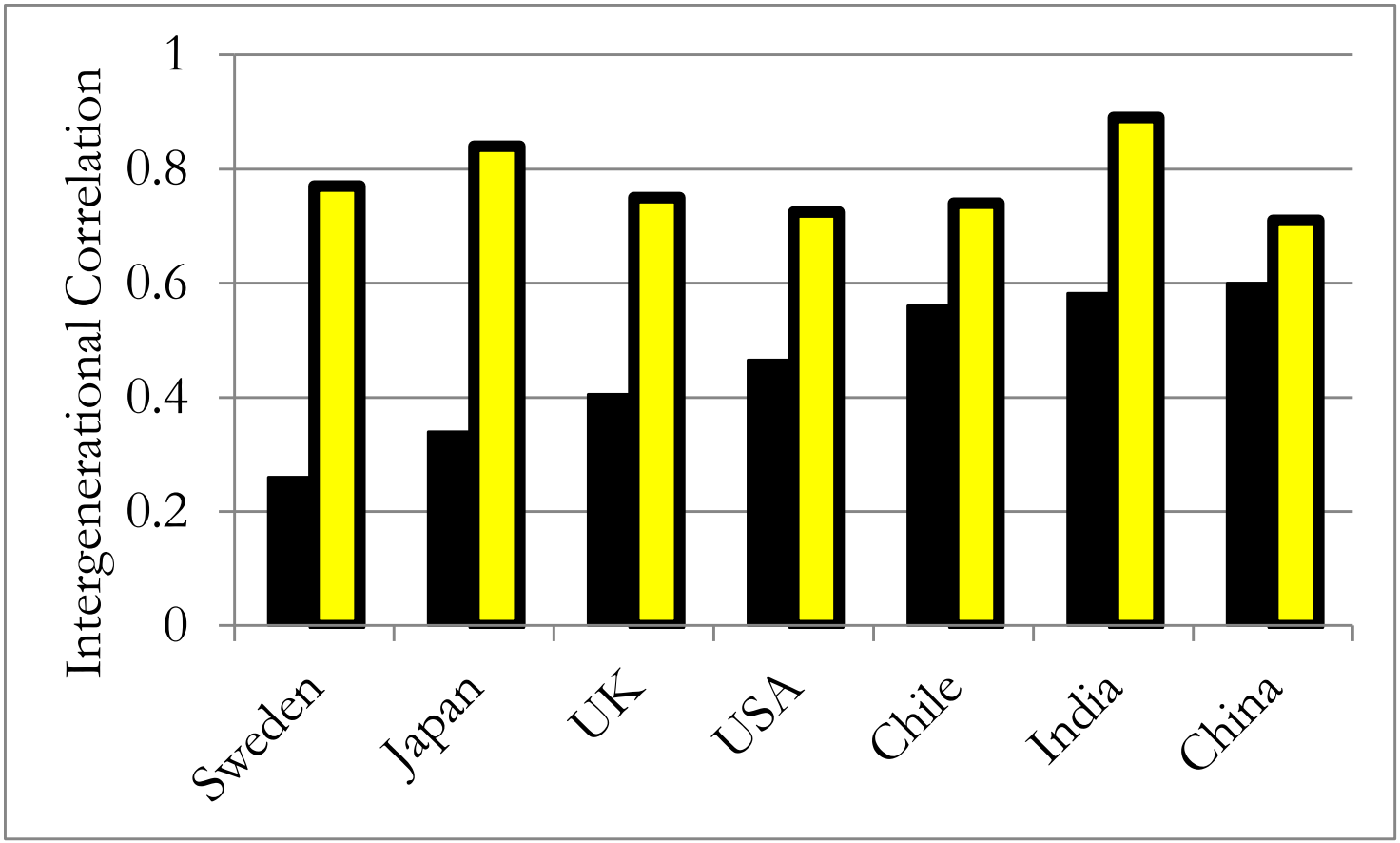
- Call this b, where k is surname grouping

$$\bar{y}_{kt+1} = b\bar{y}_{kt} + v_{kt+1}$$

- How will it relate to β?

- Both reduces and increases attenuation from errors.
- Biased by a relation between status and demography.

Conventional versus Surname Estimates of Status Persistence



Hypotheses

- b is high 0.7-0.85
- b varies little across societies and epochs
- The process of social mobility IS Markov –

$$b_n = b^n$$

- b is the same for all measures of status – wealth, earnings, education, occupation
- b is the same all across the status distribution
- Since $b^2 = 0.5\text{-}0.7$ the majority of social status, measured in this way, is determined at conception
- b is largely biologically determined

Direct and Indirect Estimates

Table 5: Rare English Surname Samples, 1858-1887

Sample A	Sample B	Sample C
Ahmuty	Aller	Agace
Allecock	Almand	Agar-Ellis
Angerstein	Angler	Aglen
Appold	Anglim	Aloof
Auriol	Annings	Alsager
Bailward	Austell	Bagnold
Basevi	Backlake	Benthall
Bazalgette	Bagwill	Berthon
Beague	Balsden	Brandram
Berens	Bantham	Brettingham
Beridge	Bawson	Brideoake
Berners	Beetchenow	Broadmead
Bigge	Bemmer	Broderip
Blegborough	Bevill	Brouncker
Blicke	Bierley	Brune

Figure 5: Log Average Wealth relative to the Average, 1858-2011

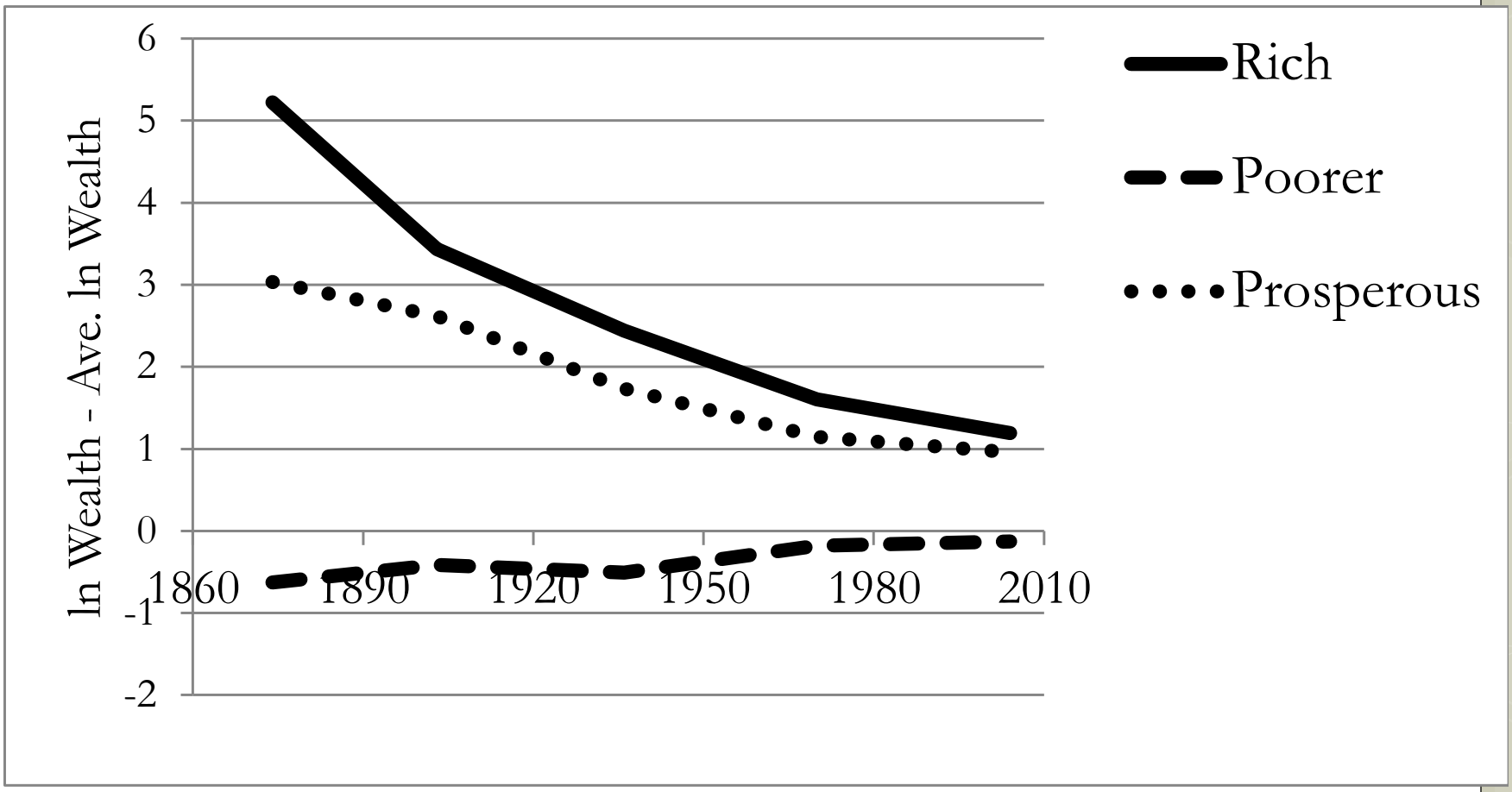


Table 6: b Values Between Death Generations

Generation	Rich	Prosperous	Rich and Prosperous	Poor
1888-1917	0.66 (0.026)	0.86 (0.052)	0.75 (0.028)	0.66 (0.061)
1918-1959	0.68 (0.031)	0.64 (0.041)	0.66 (0.030)	1.12 (0.136)
1960-1987	0.73 (0.040)	0.74 (0.051)	0.73 (0.035)	0.30 (.076)
1999-2011[a]	0.70 (0.098)	0.80 (0.125)	0.74 (0.078)	0.41 (0.615)
Average	0.69	0.76	0.72	0.61

Alternative Method of Estimating b

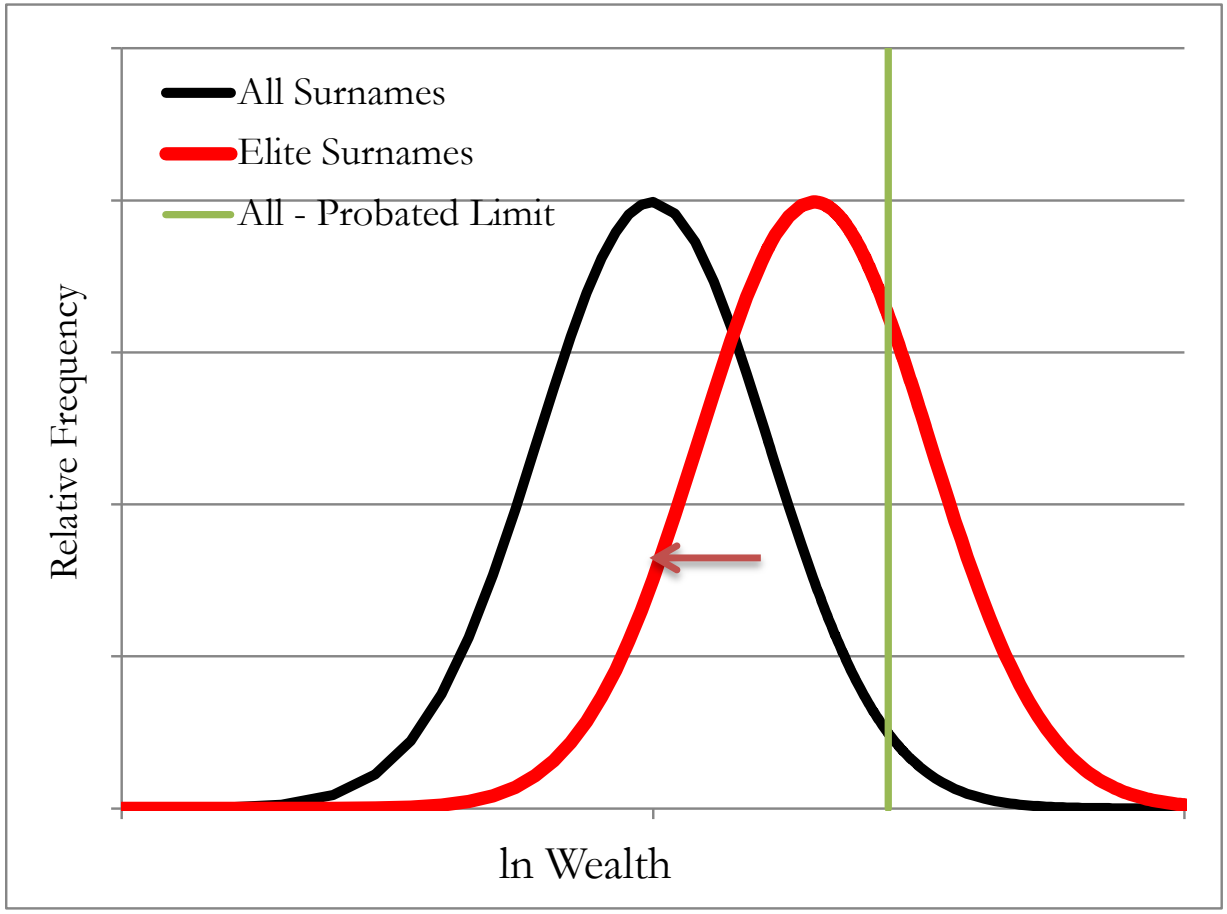
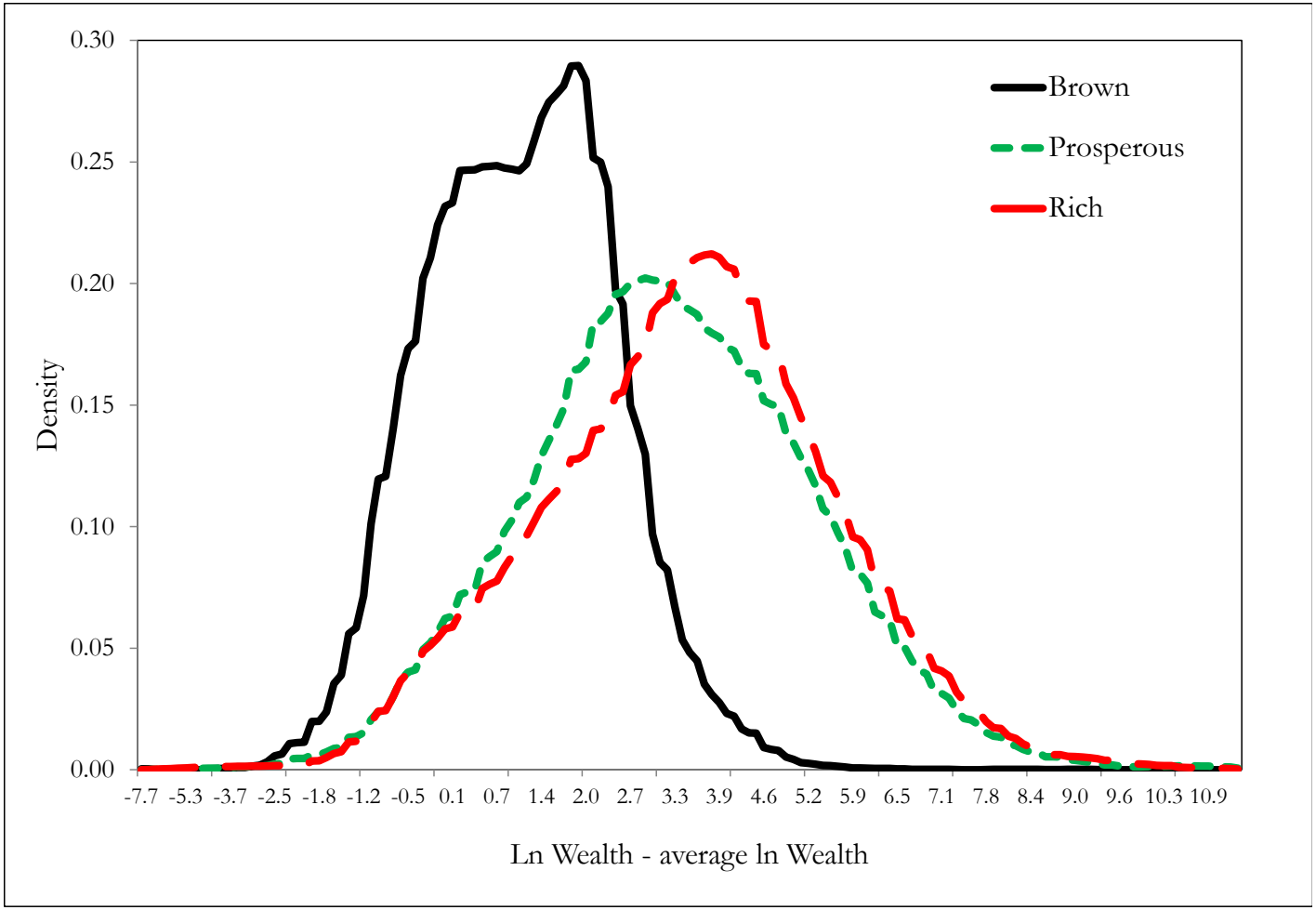


Figure 7: Wealth Distribution of Probated, Rich, Prosperous and Brown Surnames, 1918-1959



Basic Measure of Status

$$relative\ representation\ of\ z = \frac{Share\ of\ z\ in\ elite\ group}{Share\ of\ z\ in\ general\ population}$$

Table 9: Estimated b by Surname Group and Period, Probate Shares

Period	Rich 1858-87	Prosperous 1858-87	Poor 1858-87
1888-1917	0.63	0.81	0.37
1918-59	0.75	0.65	1.04
1960-93	0.59	0.70	0.80
1994-2011[a]	0.78	0.81	0.05
Average	0.69	0.75	0.57
Direct Estimate	0.69	0.76	0.61

Figure 9: Maximum Inheritance Tax Rates, UK, 1825-2012

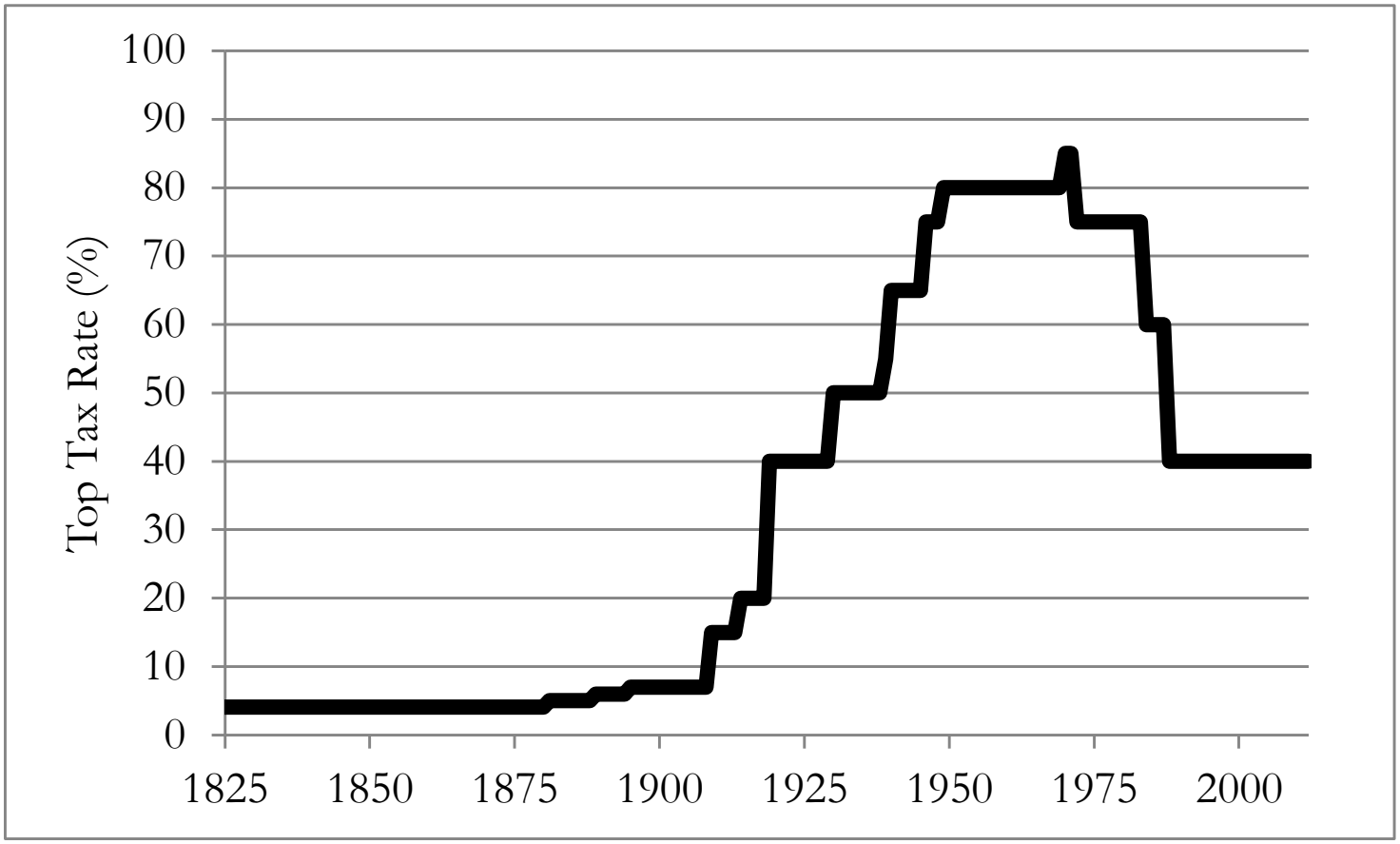


Table 9: b Estimates for England

Period	Wealth	Education	Occupations	Political Elite
1200-1400	-	0.80-0.86	-	0.91
1400-1650	0.74-0.85	0.77-0.86	-	0.91
1650-1850	0.71-0.85	0.77-0.83	-	0.91
1850-1950	0.70	0.77-0.83	-	0.81
1950-2012	0.74	0.80	0.65-1.00	-

Figure 8: The Prehistory of the Rich and Poor of 1858-87

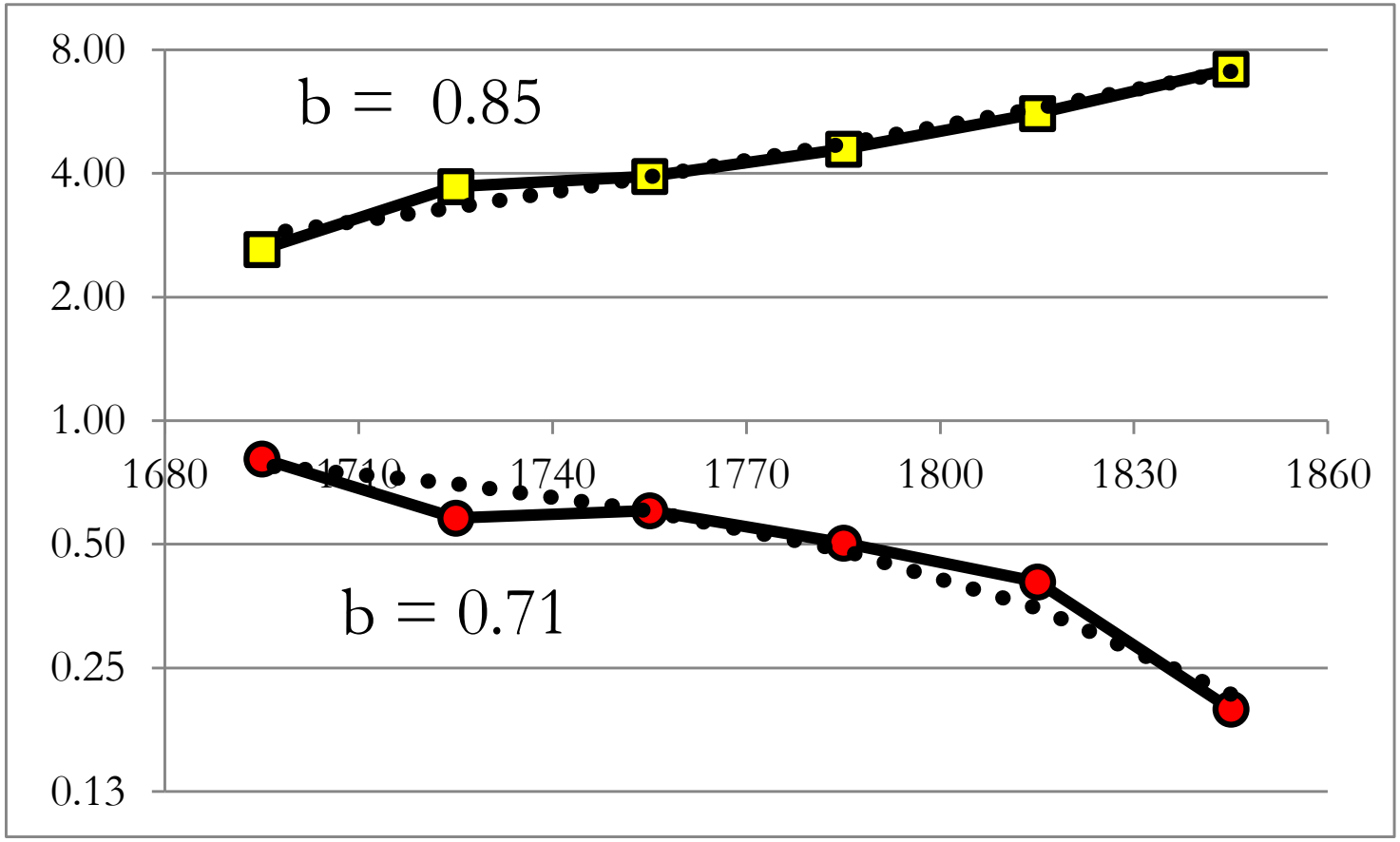
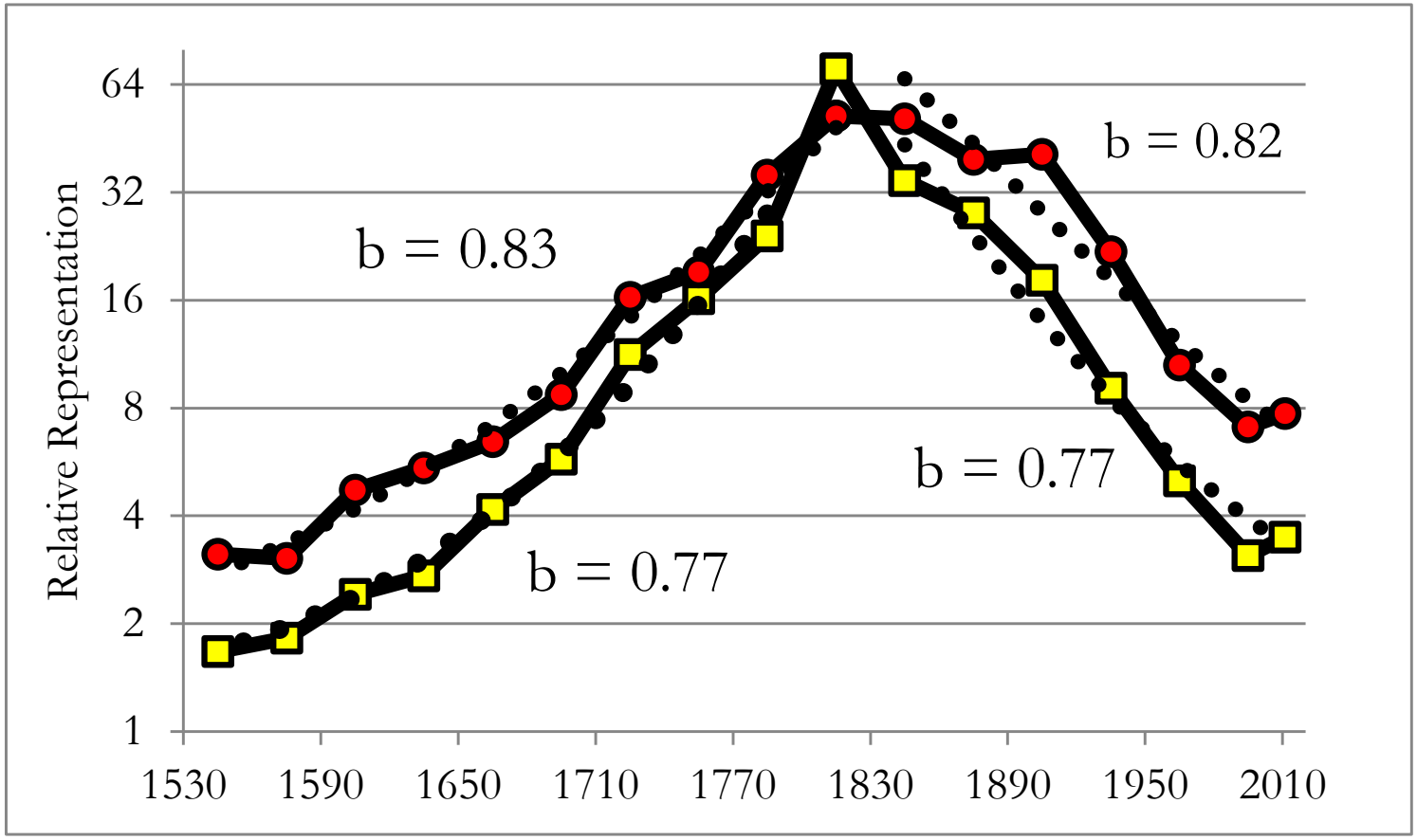


Figure 10: Relative Representation and Implied bs at Oxbridge, 1530-2012



Sweden as another example

- Elite surnames from 1600-1800
- Counts/Barons
- Untitled Nobility
- Latinized Surnames

Intergenerational Correlations in the Nordic Region

Country	b earnings	b years of education	Gini Coefficient Income
Denmark	0.15	0.30	0.25
Finland	0.18	0.33	0.27
Norway	0.17	0.35	0.26
Sweden	0.27	0.25	0.25
USA	0.47	0.46	0.41

Summary Surname b Estimates by Period, Sweden

Group	1700-1900	1890-1979	1950-2012
Attorneys	-	-	0.71
Physicians	-	0.67	0.88
University Students	0.78	0.85	0.66
Academicians	0.89	0.75	0.84

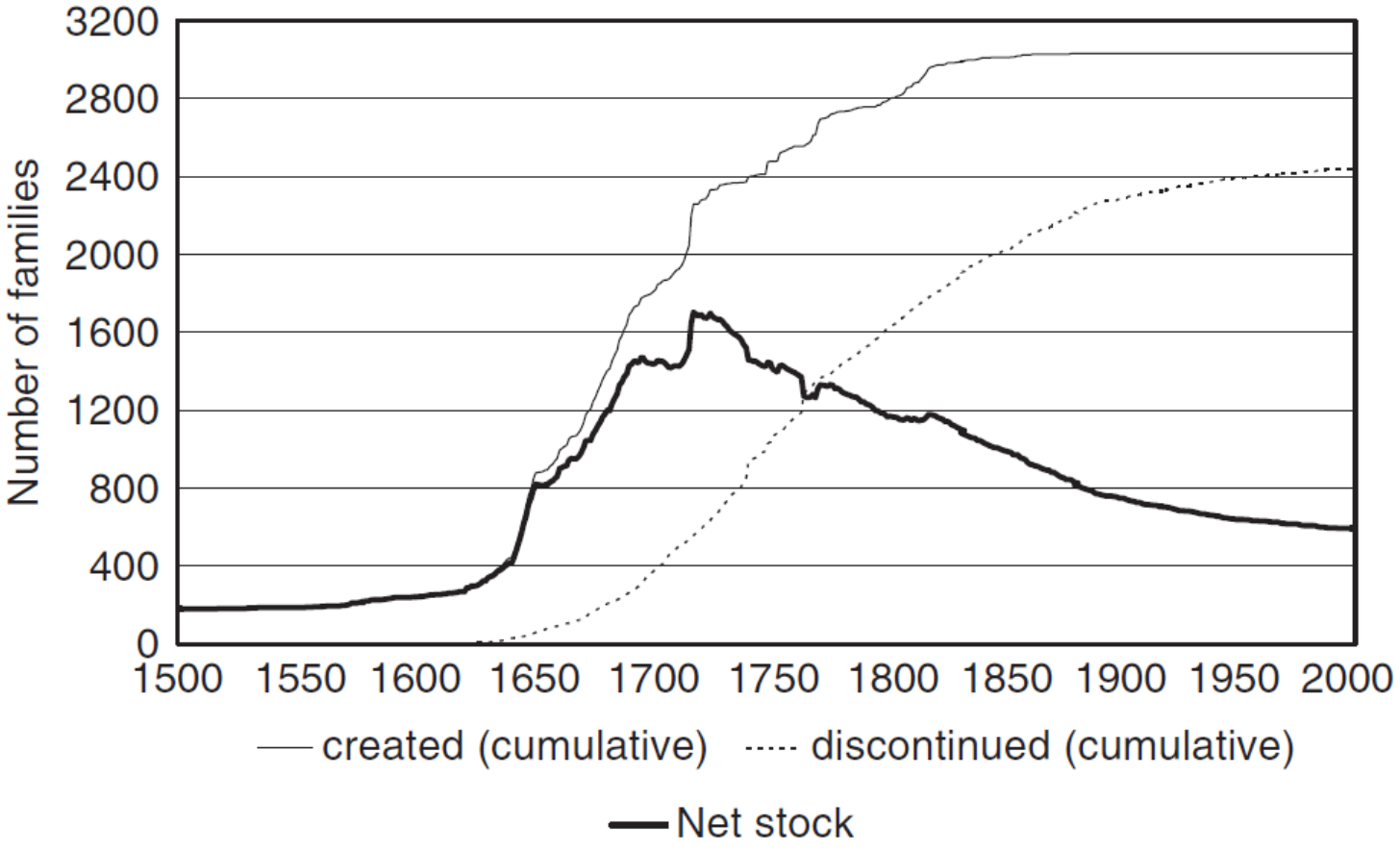
Riddarhuset, Headquarters of the Swedish Nobility

Aristocratic Surnames

- **Domestic - embodying status elements such as *Gyllen* (gold), *Silfver* (silver), *Adler* (eagle), *Leijon* (lion), and *Ehren* (honor)**

- **Leijonhufvud**
- **Gyllenstierna**
- **Oxenstierna**
- **Ehrensvärd**

Age of Aristocratic Surnames



Note: Hamlet, 1601 – Rosencrantz and Guildenstern

- *Rosencrantz* and *Gyllenstierna* were names of Danish (and Swedish) noble families of the 16th century; One tenth of the aristocrats participating in the Danish royal coronation of 1596 bore one or other name.

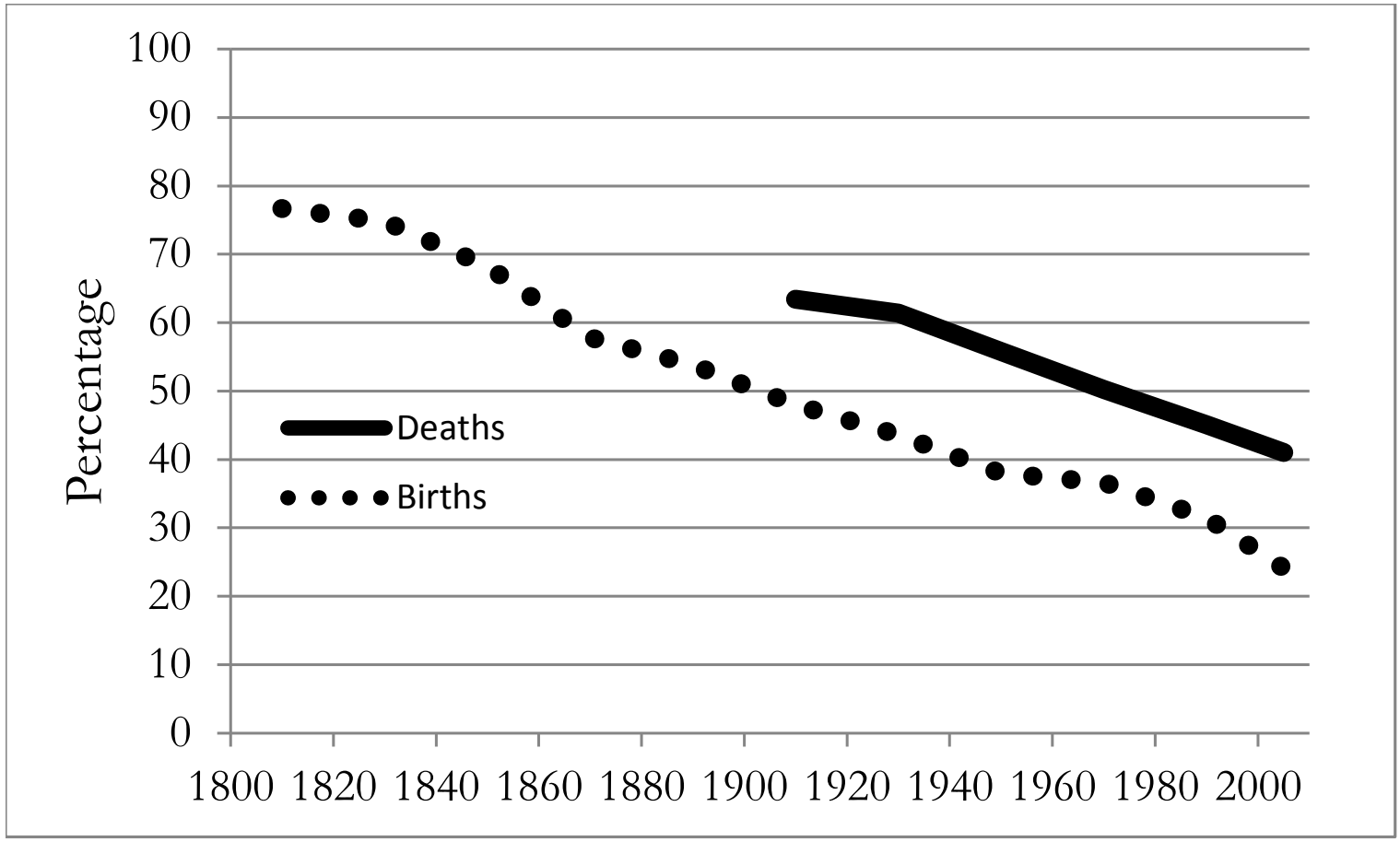
Aristocratic Surnames - Foreign

- Von Essen
- Douglas
- Bennet
- De La Gardie
- De Mortaigne

Latinized Surnames

- Celsius
- Aquilonius
- Arrhenius
- Boethius
- Bruzelius
- Cnattingius

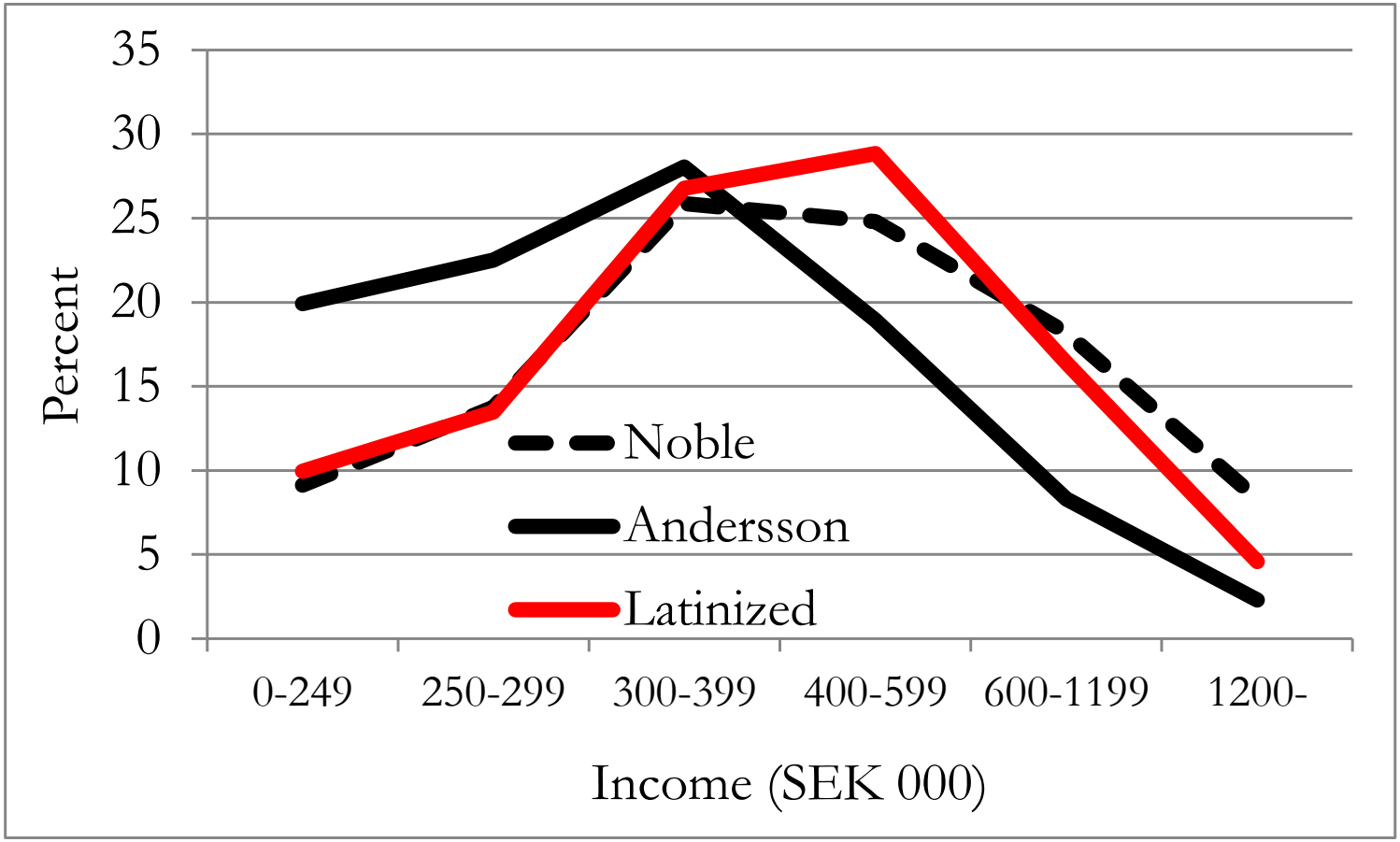
..sson surnames



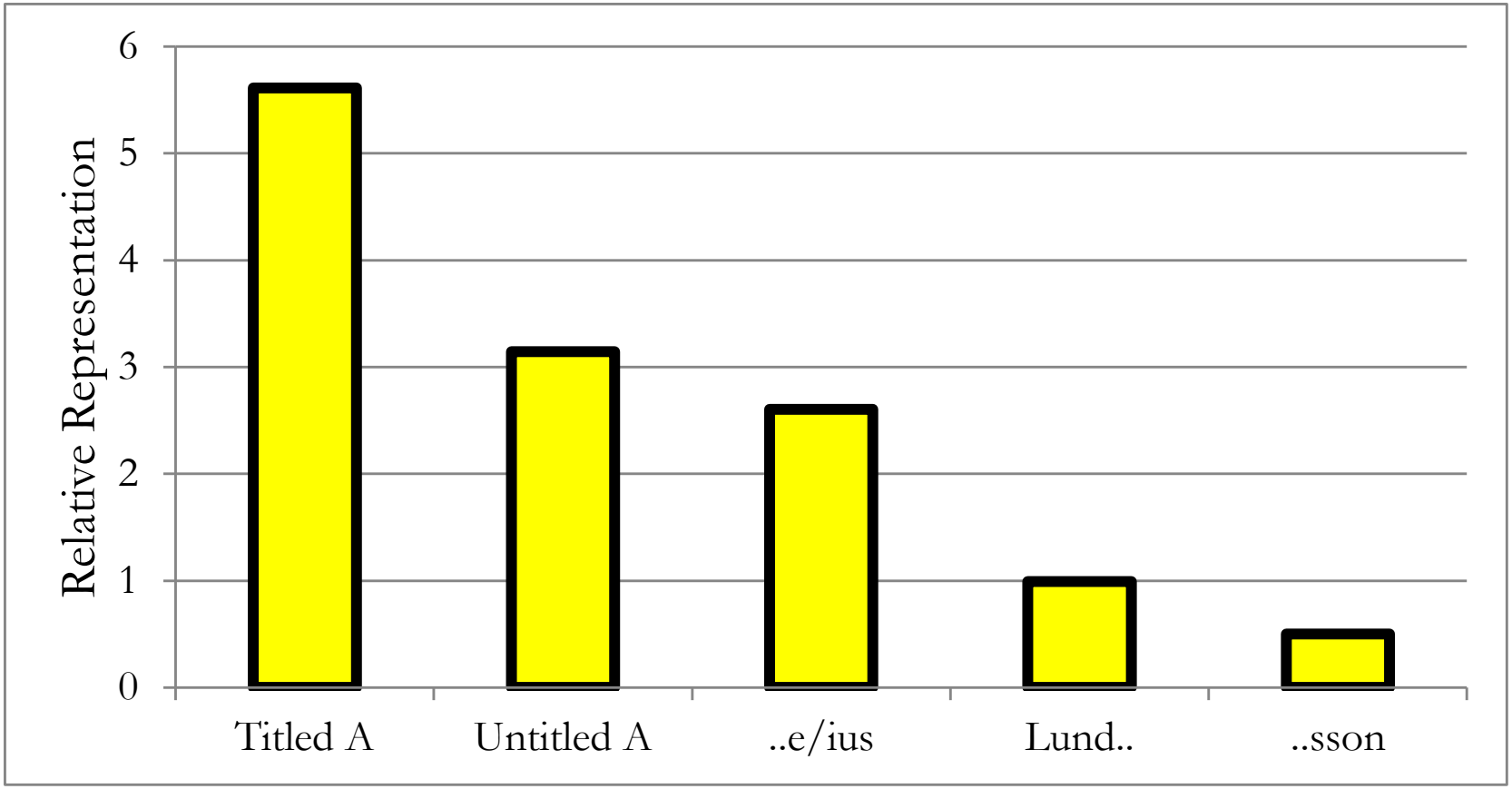
Modern Tax Information by Surname

Leijonhielm, Anna Örnbacken 26	320.400	10.131
Leijonhielm Larsson, May Backvindeln 63	283.000	
Leijonhufvud, Cecilia Banérgatan 46 2 tr	481.700	467.543
Leijonhufvud, Madeleine Basaltgränd 10	340.100	
Leijonhufvud, Margareta Bergsmarksvägen 4 1 tr	1.576.800	100.317
Leijonhufvud, Louise Blackebergsbacken 5 läg 144	119.400	1.080.423
Leijonhufvud, Eld Blanchegatan 18 4 tr	336.700	
Leijonhufvud, Margareta E C Å Hälsingehöjden 11	247.000	2.082.476
Leijonhufvud, Christina Högbergsgatan 11	279.200	
Leijonhufvud, Elisabeth Kommendörsgatan 28	573.500	
Leijonhufvud, Jenny Krukmakargatan 67 läg 0015	523.000	
Leijonhufvud, Alice Langelandsgatan 10	318.200	289
Leijonhufvud, Susanna Manhemsgatan 13 bv	283.000	
Leijonhufvud, Sven Mårdvägen 34	362.100	54.519
Leijonhufvud, Elisabet Mårdvägen 34	308.200	1.256
Leijonhufvud, Eric Nybrogatan 64	648.000	40.340
Leijonhufvud, Gustaf Nybrogatan 68 1 tr	239.500	152.518
Leijonhufvud, Titti Odengatan 23 5 tr	322.700	
Leijonhufvud, Ewa K S Ragvaldsgatan 21 4 tr	534.300	123.020
Leijonhufvud, Ruth Sigrid G Rindögatan 42	289.300	
Leijonhufvud, Fredrik Rålambsvägen 10 A	1.224.800	23.100
Leijonhufvud, Elizabeth Rålambsvägen 10 A 3 tr	667.800	

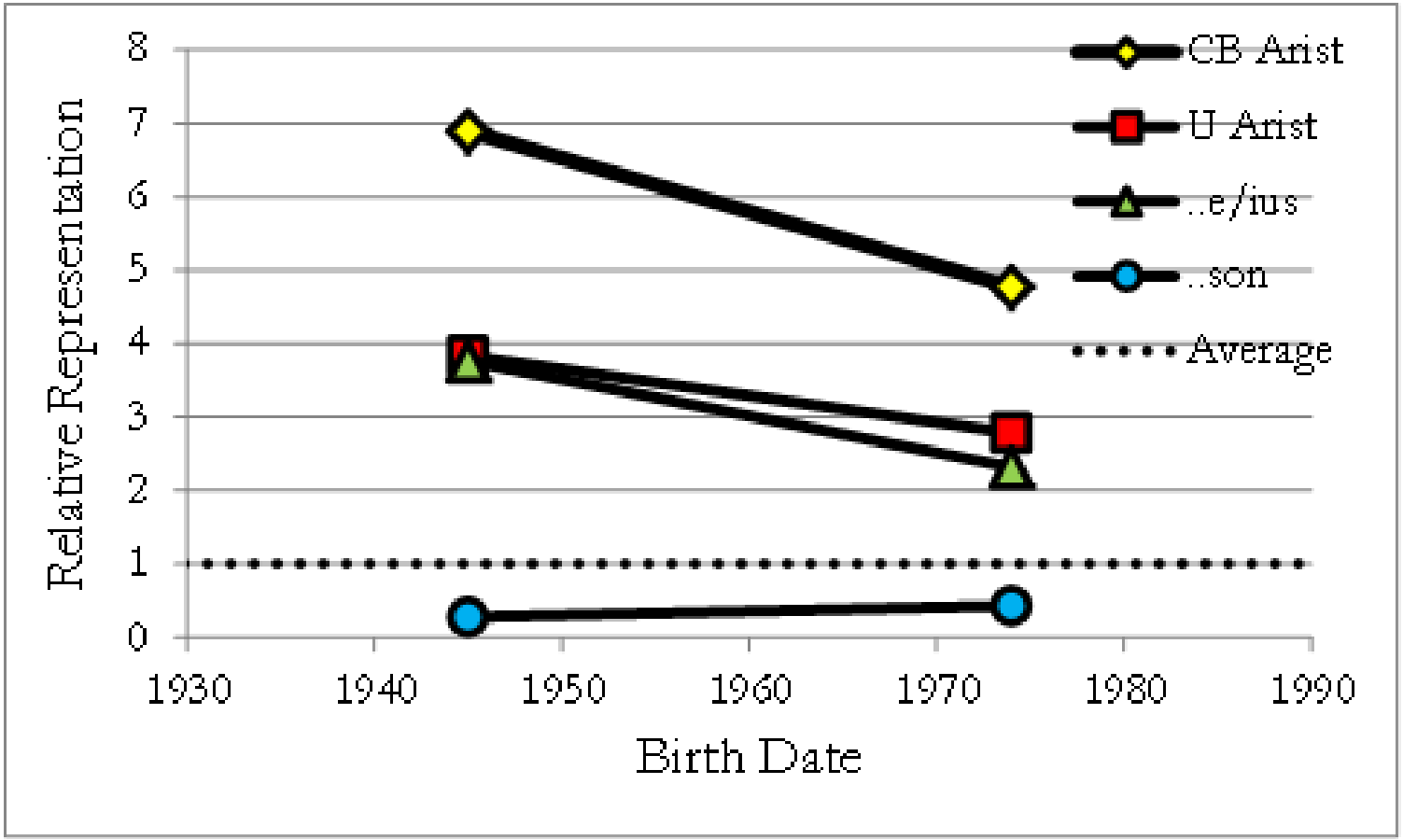
Taxable Income, 2008



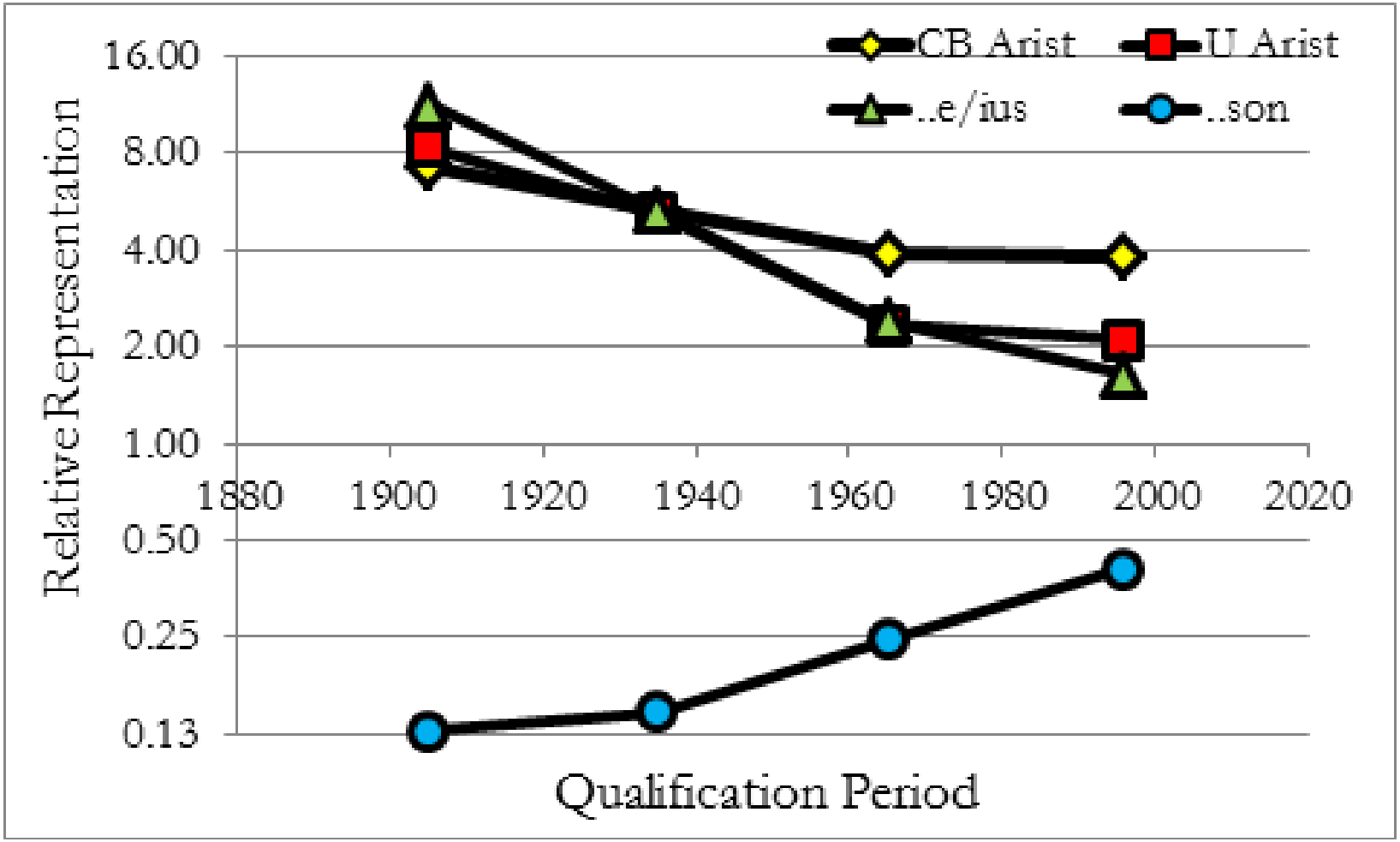
Relative Representation of Surnames, Attorneys, Sweden, 2012



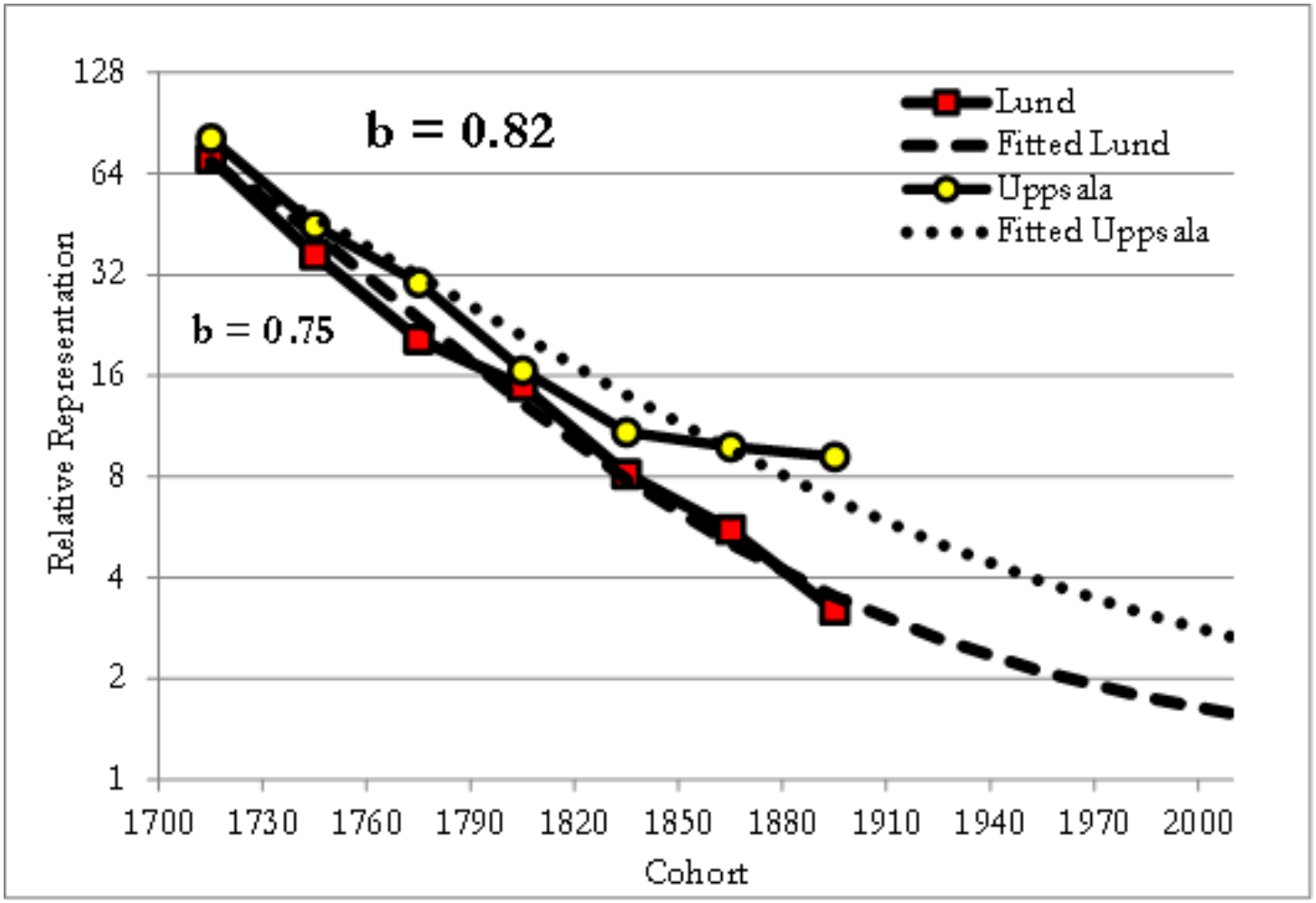
Relative Representation of Surnames, Attorneys, by Cohort



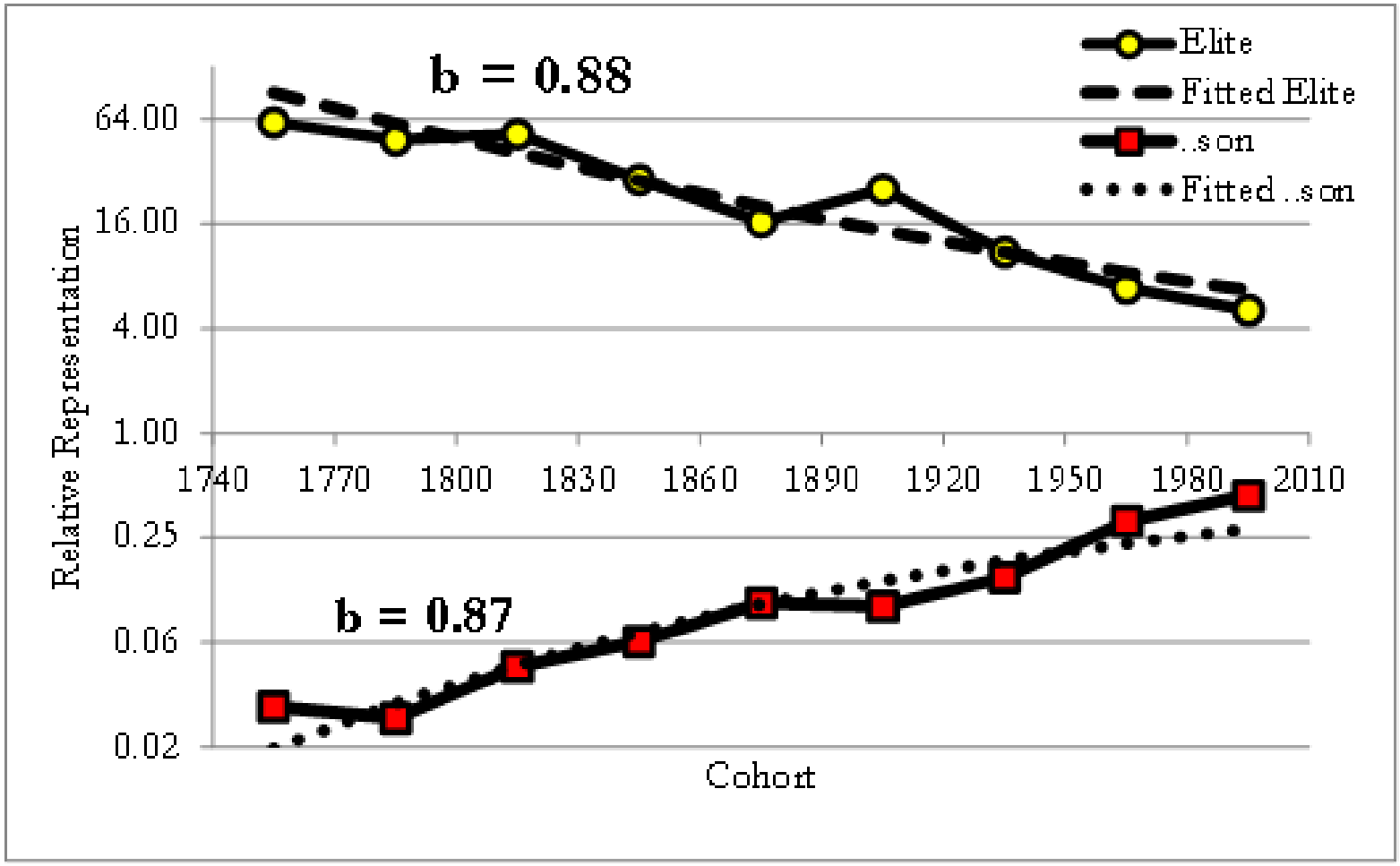
Representation of Surname Types Among Doctors, 1890-2011



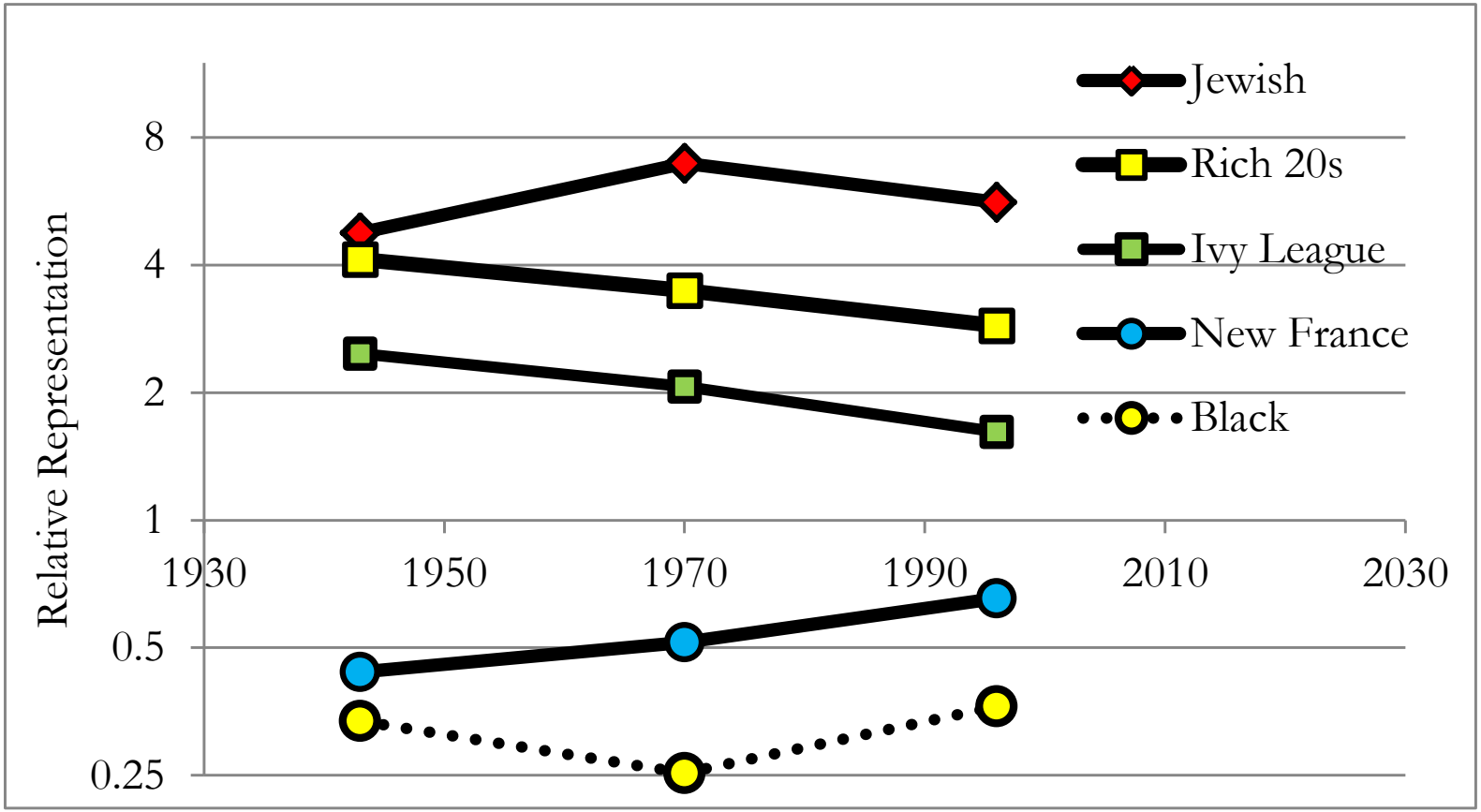
Inherited Latinized Surnames, Lund and Uppsala, 1700-1909

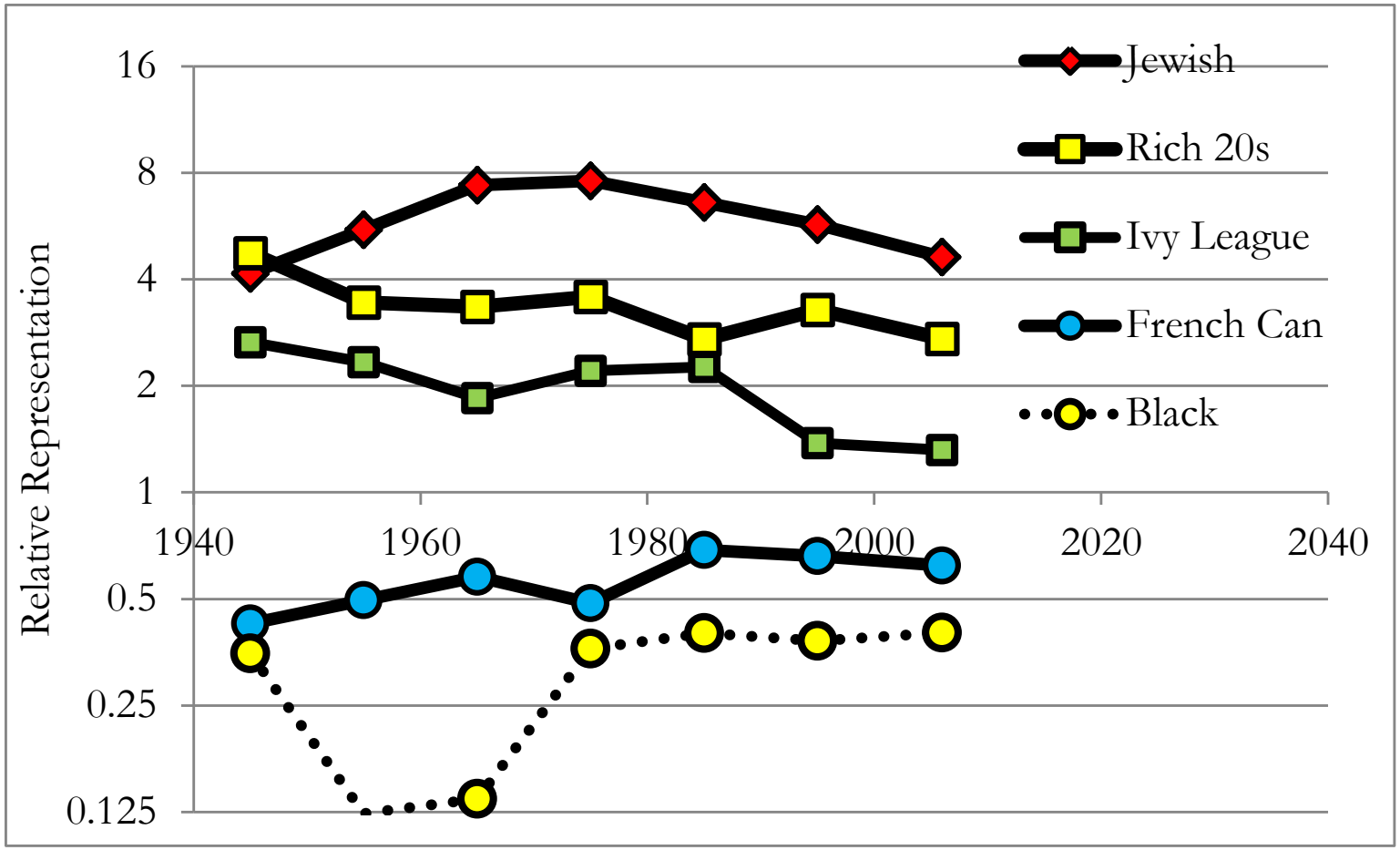


Elite Surnames in the Swedish Royal Academies



USA - Doctors


Relative Representation
8
4
2
1
0.5
0.25
1930
1950
1970
1990
2010
2030
Jewish
Rich 20s
Ivy League
New France
Black


16
8
4
2
1
0.5
0.25
0.125
Relative Representation
1940
1960
1980
2000
2020
2040
Jewish
Rich 20s
Ivy League
French Can
Black

US Doctors

Group	1920-49 to 1950-79	1950-79 to 1980-2011	1970 to 2011
Ashkenazi Jewish	-	0.88	0.75
1920s Rich	0.78	0.84	0.94
Ivy League pre 1850	0.80	0.65	0.23
New France	0.81	0.65	0.78
Black	-	0.69	0.96
Average All	0.80	0.74	0.73

US Attorneys

Group	1920-49 to 1950-79	1950-79 to 1980-2011	1970 to 2012
Katz	0.82	1.04	0.95
1920s Rich	0.84	0.86	0.95
New France	1.20	0.53	0.58
Washington	0.91	0.94	0.84
Average All	0.94	0.84	0.83

Explanation

- $y_t = x_t + e_t$
- y_t = each status manifestation
- x_t = underlying social status or competence of person

$$E(\hat{b}_y) = b_x \frac{1}{1 + \left(\frac{\sigma_e^2}{\sigma_x^2}\right)}$$

People Trade off Income, Occupation etc. in seeking Social Status

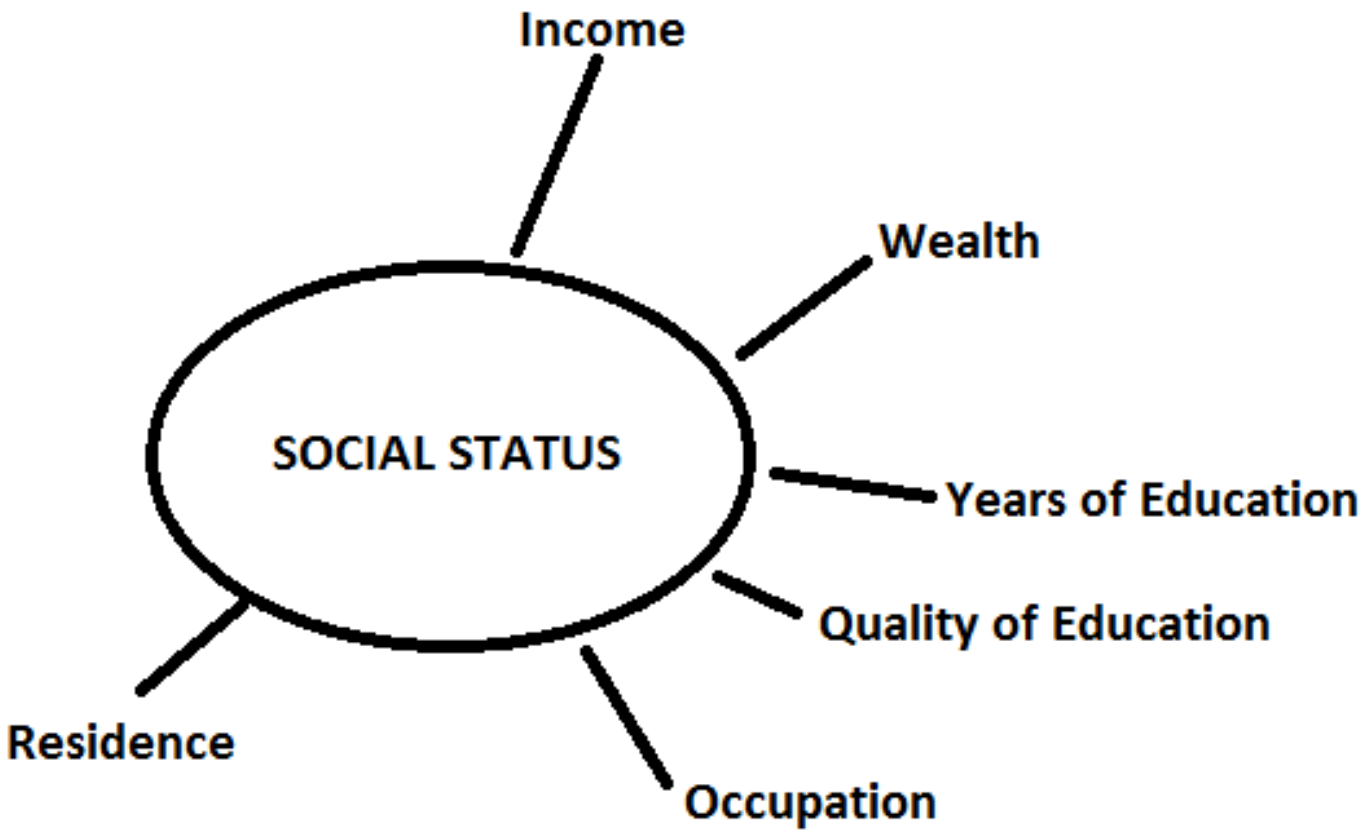


Table 2: Correlations between the Aspects of Status, Individuals

Status Element	Mental Aptitude (IQ etc.)	Education	Occupational Status	Earnings	Wealth
Mental Aptitude	-	.45-.62	.16-.31	.23-.30	.16
Education	-	-	.41-.85	.32-.34	.22-.38
Occupational status	-	-	-	.34-.71	.13-.34
Earnings	-	-	-	-	.60-.61

Implications

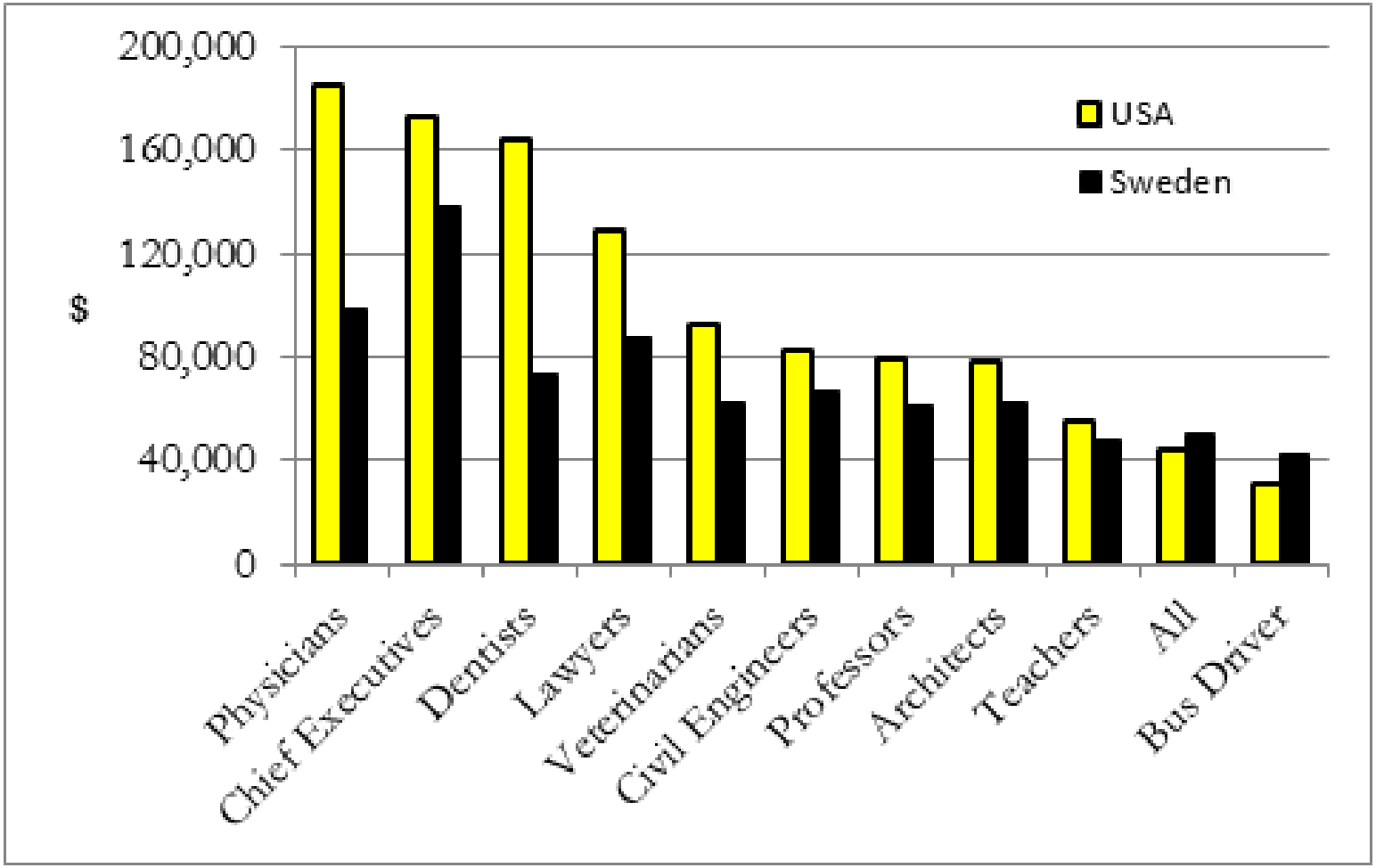
- Once we average over surnames

$$\bar{y}_i = \bar{x}$$

$$\Rightarrow b_y = b_x$$

- Assuming the underlying x has the same variance everywhere, lower inequality in y will result in a lower b_y

Average Earnings by Occupation, Sweden and the USA, 2008



Implications (cont.)

- If you took a broader measure of status – an average of earnings, education, occupation, health, wealth - $b_y \approx b_x$

- If you look at mobility for social groups not defined by status alone – Jews, Blacks, Catholics - $b_y = b_x$

- If you measure mobility over a second generation, $b_y = b_x$. Thus the process will appear non Markov, and grandparents have influence on outcomes for grandchildren

$$b_{yn} = b_x^{n-1} b_y = \varphi b_x^n$$

Implications

- If you estimate a regression with people with group differences in average x of the form

$$y_{t+1} = by_t + c_B DBLACK + c_B DJEWISH$$

- then we will find
- Cb < 0, CJ > 0

Implications

- You will also find group effects predictive of outcomes as in Borjas, 1995

- $y_{ijt+1} = b_0 y_{ijt} + b_1 \bar{y}_{jt}$

- But not because of "ethnic capital"

Table 3: Persistence in Education Across Multiple Generations in Sweden

Last Generation	Great-Grandparents	Grandparents	Parents
OBSERVED			
Grandparents	0.334		
Parents	0.229	0.312	
Children	0.123	0.202	0.412
PREDICTED, b = 0.7			
Grandparents	0.334		
Parents	**0.226**	0.312	
Children	**0.173**	**0.253**	0.412

Source: Lindahl et al., 2012, table 2.

More Fundamental Question – what transmits social genotype?

- Human Capital? (Investment)
- Genes?
- Culture?

Tests? Human Capital

- Intergenerational correlation greater for capital constrained as opposed to unconstrained
- Why hasn't b dropped in recent years?
- Why isn't b higher for lowest income groups?

Tests? Human Capital

- Solon (2012), elaboration of Becker and Tomes, 1979

- $b_y = \frac{(\gamma+\tau)}{(1+\gamma\tau)}$

- τ is the correlation of abilities
- γ is the elasticity of income with respect to human capital investment

Tests? Genes versus Culture

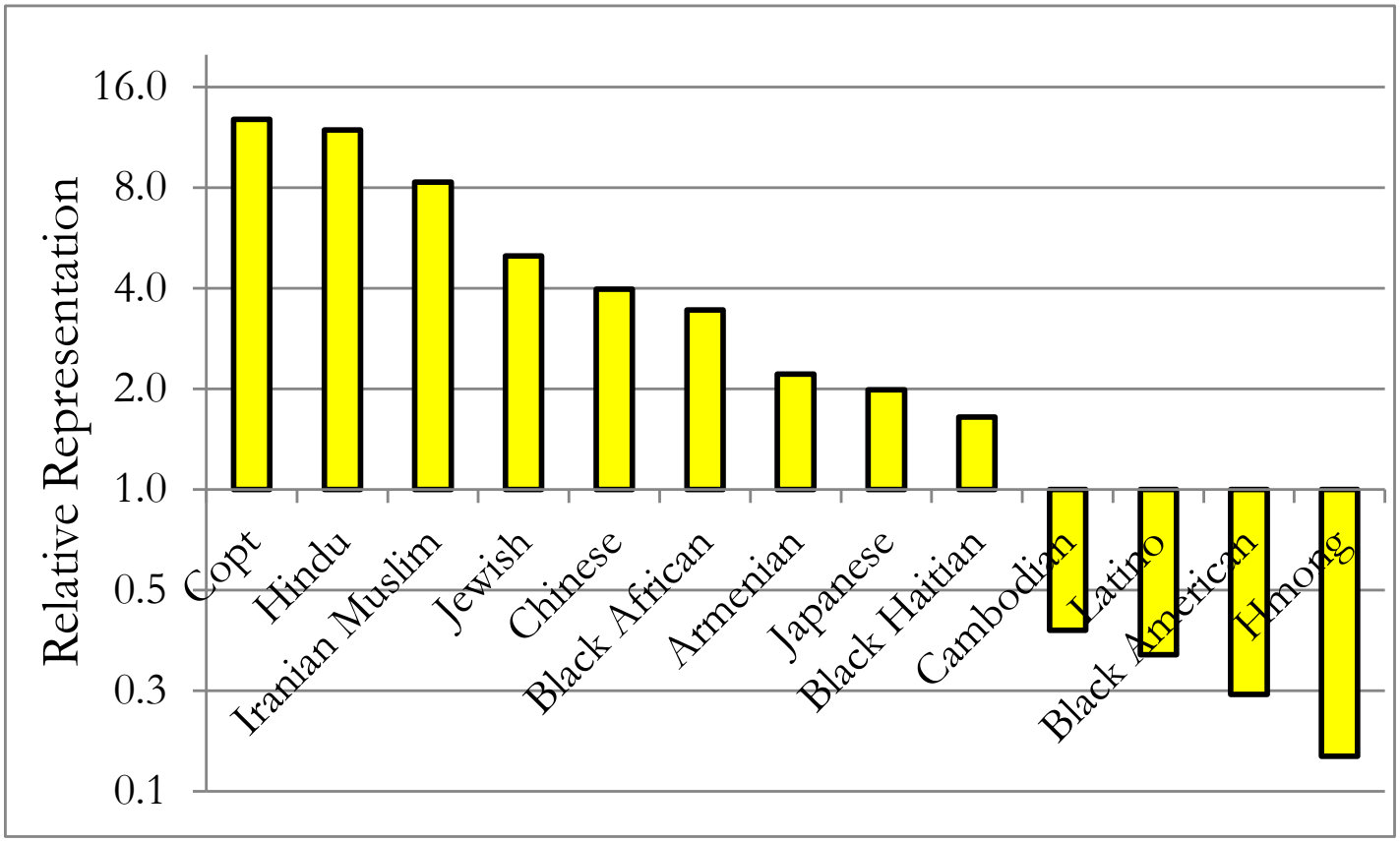
- Adoption Studies – Sweden, US
- Genes contribute 3-4 times as much to variance of outcomes as family environment.
- With an interaction between genes and environment the ultimate effects of genes could be greater.

Tests? Genes versus Culture

- Are elites and underclasses all just selective draws from the population distribution? Or do such groups have distinctive cultures?

- Endogamy – with genes as the drivers of status this should stop group regression to the mean.

Figure 26: Frequency of Doctors across Immigrant and Domestic Surname Groups, USA, 2012



Note: The vertical scale measures the numbers of doctors per capita in these groups relative to the population of the USA as a whole.

Persistent Elites

- Egyptian Copts, Jews 700-2012
- Eastern European Jews 1300-1945
- Christians in Syria, Jordan, Iraq, Iran 700-2012
- Brahmins in India