Surnames and the Laws of Social Mobility

Gregory Clark, University of California, Davis
gclark@ucdavis.edu

What is the true rate of social mobility? Modern one-generation studies suggest considerable regression to the mean for all measures of status – wealth, income, occupation and education across a variety of societies. The $\beta$ that links status across generations is in the order of 0.2-0.5. In that case inherited surnames will quickly lose any information about social status. Using surnames this paper looks at social mobility rates across many generations in England 1086-2011, Sweden, 1700-2011, the USA 1650-2011, India, 1870-2011, Japan, 1870-2011, and China and Taiwan 1700-2011. The underlying $\beta$ for long-run social mobility is around 0.75, and is remarkably similar across societies and epochs. This implies that complete regression to the mean for elites takes 15 or more generations.

NOTE – THIS VERSION STILL INCOMPLETE

Introduction

Using surnames this paper examines social mobility rates over many generations, and across countries, time periods, and different measures of social status. The framework adopted is very simple. I assume that we have measures of status that are cardinal, or can be approximated as cardinals: income, wealth, years of education, level of education, or occupational status. Then if $y_i$ is the logarithm of such a measure of status, the intergenerational elasticity, $\beta$, is inferred just as the regression coefficient from

$$y_{i+1} = \beta y_i + u_i$$

(where the mean of $y_i$ has been normalized to 0). $1 - \beta$ is the rate of regression to mean. If there is a constant variance across generations of this measure of status, then $\beta^2$ is share of social position variance that derives from inheritance. Assuming also that the process of transmission of status is AR1, then $\beta^o$ is the intergenerational elasticity of status over n generations.
There have been over the last 40 years many measures of the intergenerational elasticity of various measures of status within this framework. Figure 1, for example, shows estimates of the intergenerational elasticity of income for a variety of countries summarized by Corak (2011).

These studies suggest the following conclusions:

(1) Intergenerational elasticities are typically of the order of 0.2-0.5 for income, years of education, occupational status, and even for wealth.

(2) This implies, assuming that $y_{t+n} = \beta^0 y_t + u_{t+n}$, that income, occupational and social mobility are all largely complete within 2-5 generations. Thus a person in generation 0, with an income 20 times above the average, or 1/20 of the average, will have descendants 5 generations later with on average incomes within 10% of the average.
(3) The fraction of variance of social position explained by inheritance is low. The above figures suggest this is 4% in Scandinavia, and 22% in the USA (assuming the variance of log income is stable across generations). Most of social status is not predictable at birth.

(4) Social mobility rates vary substantially across countries.

(5) Thus the mobility rates are “too low” in some societies. With better opportunities for the children of low income or status families, more mobility would be possible.

However, as a measure of the intergenerational transmission of social status, these one generation studies suffer a key limitation. Suppose in particular we assume that the various aspects of social status in generation t – income, wealth, education, location – are all linked to some fundamental social competence or status, x_t, such that

\[ y_t = x_t + \epsilon_t, \]

where \( \epsilon_t \) is some random component. Then the regression to the mean exhibited by each partial measure of underlying status \( y \) will overestimate the regression to the mean of that underlying status. When we classify families as high or low status based on partial measures such as income, wealth, education or occupation there will appear to be substantial regression to the mean. But if we took a more aggregate measure of status, the regression will be substantially lower. And even that aggregate status regression calculated for a first generation t to t+1 will be greater than that observed in the same families from generation t+1 to t+2 and so on.

So the intergenerational elasticity estimated from two generation studies will greatly overestimate social mobility in the long run for two reasons: it measures the regression just of particular aspects of status, and it incorporates a regression through those of measured high status having net positive random components in status only in the first generation, which will not occur across future generations. The greater are the random components in determining measures of status such as income, the greater will be the degree of mismatch between such partial one generation estimates of regression to the mean and the underlying regression of fundamental social status.

---

1 We can show evidence of this for England where wealth, educational status, and occupational status are closely linked for surname cohorts.
I can thus easily construct a simple example in this framework where people are of high or low status, which is inherited perfectly,

\[ x_{i,t+1} = x_i \]

but where because observed log income \( y_i = x_i + \epsilon_i \) the \( \beta \) estimated by regressing log income in one generation on the next is 0.5. This is shown in figure 2 where the high status group (the red triangles) have an inherited fundamental log income of 5, compared to the low status group (the yellow squares) with an inherited fundamental log income of 3. There is effectively a class society in this example, where after the first generation there will be no further regression to the mean of those of high income. To estimate the rate of mobility for underlying social status we need a different method.

I show below that by using surname cohorts we can estimate the underlying \( \beta_s \) where

\[ \beta_y = \beta_s \theta, \quad \theta < 1 \]

using just the same information on the observed social status measures \( y \) which cause unknown underestimation in the two generation studies.
**Surnames**

To investigate the rate of regression to the mean of this deeper underlying social status (and by implication the long run rate of regression to the mean of income, wealth, occupational status and education) this study traces people not through individual family linkages, but through surnames over multiple generations.

In many societies surnames are inherited unchanged from one generation to the next, typically through the patriline. If at some generation surnames differ in social status, we can then trace through surnames the descendants of the current generation for many generations. As long as there is nothing peculiar about the path of descent of surnames, the surnames link the status of groups of families many generations in the past with their descendants in the present.

When initially formed, surnames in many societies were associated with social status. For example, in England some high status land owners already possessed surnames at the time of the Domesday Book of 1086, which listed the major landholders of England. Most of these people were the Norman, Breton and Flemish followers of Duke William of Normandy, who seized the throne of England in 1066. These surnames thus constitute a distinctive subset of modern English surnames: Baskerville, Beaumont, D'Arcy, de Vere, Mandeville, Montgomery, Vernon, and Villiers, for example. In England also about 10 percent of surnames derive from the occupations of the original holder, and these occupations had a range of social statuses: Smith, Baker, Shepherd, Clark, Chamberlain, Butler.

In Sweden, surnames started as patronyms which changed with each generation. Sven, son of Lars, was Sven Larsson. But his son Gunnar would be Gunnar Svensson. For the ordinary people patronyms did not become fixed across generations until the late nineteenth century. However, from at least the 17th century two groups of high status individuals were acquiring permanent and distinctive surnames. The first were those who attended universities, who adopted latinized or grecified surnames such as Celsius, Linnaeus, and Melander. The second was the aristocracy, often imported mercenary commanders, who imported surnames from Germany, Scotland and elsewhere or created their own distinctive family names when inducted into the house of nobles such as Leijonhufvud.
Even in societies such as England where the early introduction of universal surnames by 1300 meant that by 1800 common surnames all had the same average social status, we can study modern long run social mobility through the use of rare surnames. Through processes of chance in each generation some such rare surnames will be on average of high status, others of low status. If in some initial generation, surnames are predictive of social status, then over time, as long as $\beta$ is less than 1, surnames should lose this information. And the rate at which they lose it is a measure of the rate of social mobility. If the high rates of mobility typically found in one generation studies are predictive of long-run rates of social mobility, then within a few generations surnames should contain no systematic information on social status.

The crucial advantage the surname linkages give is that I can identify high and low status groups in some initial period, and then track them over multiple generations using their initial classification of status into high and low groups. This means that after the first generation the average error from the underlying status associated with each surname group in each generation is 0, so that for the surname cohorts

$$\beta_y = \beta_s$$

The $\beta_s$ estimated for surnames, however, is not identical to that within families, if we could estimate that. This is because in surname cohorts, when we estimate

$$\tilde{y}_{kt+1} = a + \beta\tilde{y}_{kt} + u_{kt+1}$$

$\tilde{y}_{kt}$ measures, for example, the average log wealth across a group of people with the surname k in the initial generation. But some of these people will not have any children, and would not be included in the within family regression. And those with 1 child from generation t get weighted as much as those with 10 children. Thus surname cohorts themselves introduce some measurement error in $y_n$ which will reduce the observed value of $\beta$. The magnitude of this downwards bias will decline, however, the larger the size of the surname groupings unless there is some systematic connection between social status and child numbers.

There will, however, be such systematic connections in some periods. In England, for example, between 1200 and 1800 high status individuals had more children, while between 1870 and 1950 they had fewer children. For 1870-1950 in
England thus, the surname method will tend to overweight the high status in the initial period, and thus underestimate the true $\beta$. 1200-1800 the estimates will instead overweight the poorer individuals in any surname group, and will thus overestimate the size of the true $\beta$. However, I observe empirically below that this bias is modest.

In looking at social mobility through surnames in some cases I have direct measures such as wealth in England 1858-2012. However, in most cases I have instead measures of the fraction of people bearing a surname who are in high or low status occupations over many generations compared to the fraction of those surnames in the general population: university graduates, doctor, attorney, member of Parliament, professor, author, or criminal. What I observe over generations is the relative representation in some top percentile $x$ of social status of various surnames, defined as

$$relative\ representation = RR = \frac{surname\ share\ of\ group\ x}{surname\ share\ of\ population}$$

To extract implied $\beta$s for these cases I proceed as follows. Assume that social status, $y$, follows a normal distribution, with mean 0 and variance $\sigma^2$. Suppose that we know the initial status of a surname, $z$, within the social status distribution. We might know, for example, that the occupational surname “mason” referred to an occupation in the middle of the status distribution. Its relative representation among high social status groups, such as university graduates would start at 0. However, while the variance in status for this surname starts at 0, the variance after $t$ generations will be

$$\sigma_{zt}^2 = (1 - \beta^{2t})\sigma^2$$

(2)

Thus the time for the surname to achieve a relative representation of 1 amongst various elites will provide estimates of the average level of $\beta$. Indeed even if the initial status of the name deviates from the mean, as long as the location of the surname is not too far from mean status, the important element in limiting the approach of this surname status distribution to the general distribution will be the rise of the variance to that of the society as a whole. Also the path of relative representation for generations 1, $\ldots$, $t$, $\ldots$ will show whether $\beta$ can be approximated as a constant for all generations.

Another situation that arises is that a surname, $z$, has a relative representation greater than 1 among elite groups. The situation looks as in figure 3, which shows the general
pdf for status (assumed normally distributed) as well as the pdf for the elite group. The problem here is that a given relative representation is consistent with a range of initial values for the mean and variance of social status for this surname \( \overline{y}_{z0}, \sigma_{z0}^2 \).

But for any assumption about \( \overline{y}_{z0}, \sigma_{z0}^2 \) there will be an implied path of relative representation of the surname over generations for each possible \( \beta \). This is because

\[
\overline{y}_{zt} = \overline{y}_{z0} \beta^t
\]  
(3)

Also since \( \text{var}(y_{zt}) = \beta^2 \text{var}(y_{zt-1}) + (1 - \beta^2)\sigma^2, \)

\[
\text{var}(y_{zt}) = \beta^{2t} \sigma_{z0}^2 + (1 - \beta^{2t})\sigma^2
\]  
(4)

With each generation, depending on \( \beta \), the mean status of the elite surname will regress towards the population mean, and its variance increase to the population variance (assuming that \( \sigma_{z0}^2 < \sigma^2 \)). Its relative representation in the elite will decline in a particular pattern.

Thus even though we cannot initially fix \( \overline{y}_{z0} \) and \( \sigma_{z0}^2 \) for the elite surname just by observing its overrepresentation among an elite in the first period, we can fix these by choosing them along with \( \beta \) to best fit the relative representation of the elite surname \( z \) in the social elite in each subsequent generation.
While we can in general expect that

\[ 0 < \sigma_{z0}^2 < \sigma^2 \]

it turns out to matter little to the estimated size of \( \beta \) what specific variance is assumed. Consider the case, as in figure 3, where the majority of the high status surname group still lies outside the observed elite. If we assume \( \sigma_{z0}^2 = \sigma^2 \) then for a given \( \beta \) we will have the quickest convergence on the population distribution, since the variance of this surname’s status is already at the population average, and the implied initial average status of this surname, \( \bar{y}_{z0} \) will be closest to the population mean. In contrast the case in which for a given \( \beta \) the elite would take the longest time to be distributed as is the general population is that where \( \sigma_{z0}^2 = 0 \), and the mean status of the elite group is exactly at the upper 2% level of the distribution. So for any length of time \( T \) until effective convergence we can easily find the upper and lower bound implied for \( \beta \).

Suppose for example that the relative representation of an elite in the top 2% of the status distribution is 8, and that it takes 10 generations for that relative representation to fall below 1.1. If status is normally distributed both among the general population, and among the elite, what is the possible range of \( \beta \)? The answer is that \( \beta \) lies between 0.65 and 0.70: 0.65 if the initial variance of the elite status was 0, 0.70 if their initial variance was the same as that for the population. Thus if we assume that the status variance of any elite is the same as that of the population as a whole we will get an upper bound estimates of the level of \( \beta \) by observing time to convergence.

Figure 4 shows the path of the relative representation of this group in the top 2% under each assumption, and the constraint that the relative representation falls below 1.1 by the tenth generation. Note that under the assumption of initially no variance among the elite, so that they are all clustered exactly at the upper 2% boundary, in the initial generation their share about 2% actually increases, since the fall of their mean is counteracted by increased variance.

---

2 Where a majority of a group lies above the observed threshold, the assumption that would provide quickest convergence for a given \( b \) would be 0 variance in the group, since that would produce the lowest group mean status.

3 The assumption that produces the highest \( \beta \) is zero variance initially if a majority of the surname group is in the elite initially.
Where we observe the relative representation of an elite generation by generation, we can use the different shaped convergence paths that different assumptions about initial variance imply to make an estimate of the initial variance also, and hence a more precise estimate of $\beta$.

**Figure 4: Convergence Path, Different Assumptions about Initial Variance**

![Convergence Path](image)

To illustrate how this estimate works in practice consider the data in table 1. This shows the relative representation at Oxford and Cambridge Universities in England of high average wealth rare surnames, based on the wealth at death of those born 1780-1809 who died 1858 and later. In 1800-1829 the high wealth surnames show up at 94 times their share in the population among entrants to Oxford and Cambridge. Relative representation for this elite group declines very little in the years 1830-59, for the children of the first generation. We thus take this second generation as the baseline, and ask what the subsequent decline implies about the rate of social mobility.
Table 1: Relative Representation of Rare Surnames at Oxbridge, 1800-2010

<table>
<thead>
<tr>
<th>Period</th>
<th>Sample Size</th>
<th>N Wealthy Surnames</th>
<th>Relative Representation Wealthy Surnames</th>
<th>Relative Representation Any Rare Surnames 1800-29</th>
</tr>
</thead>
<tbody>
<tr>
<td>1800-29</td>
<td>18,651</td>
<td>169</td>
<td>94</td>
<td>117</td>
</tr>
<tr>
<td>1830-59</td>
<td>24,418</td>
<td>210</td>
<td>91</td>
<td>49</td>
</tr>
<tr>
<td>1860-89</td>
<td>35,503</td>
<td>184</td>
<td>55</td>
<td>34</td>
</tr>
<tr>
<td>1890-1919</td>
<td>22,005</td>
<td>77</td>
<td>43</td>
<td>19</td>
</tr>
<tr>
<td>1920-49</td>
<td>44,231</td>
<td>73</td>
<td>25</td>
<td>9.8</td>
</tr>
<tr>
<td>1950-79</td>
<td>95,792</td>
<td>67</td>
<td>9.1</td>
<td>6.3</td>
</tr>
<tr>
<td>1980-2010</td>
<td>213,303</td>
<td>65</td>
<td>9.2</td>
<td>4.0</td>
</tr>
</tbody>
</table>

Figure 5: Relative Representation at Oxbridge, 1830-2010

\[ b = 0.79 \]

\[ b = 0.76 \]
The table shows that the rich rare surnames steadily converging in relative representation towards 1. However, the rate of convergence is slow. Even for the cohort entering Oxbridge 1980-2010 the rich rare surnames are still 9 times more frequent relative to the stock of 18 year olds with that name than are common indigenous English names such as Brown(e) or Clark(e).

What does the pattern in decline of relative representation shown in table 1 imply about the $\beta$ for education? If we assume a normal distribution of status, and that all those of high status had the same variance as the general population, then we can estimate what the $\beta$ for educational status 1830-2010 was. Since the high status surnames had a relative representation of 91 among the top 0.7% of the educational hierarchy in 1830-59, this fixes what the mean status of those names had to be, relative to the social mean, assuming the variance of their status was the same as that of the general population. For each possible $\beta$ their relative representation would decline generation by generation in a predictable manner. Figure 5 shows the actual pattern, as well as the single $\beta$ that best fits the data. For the wealthy group that is $\beta = 0.79$. Notice also that there is no sign that educational mobility has speeded up in the last few generations. The single $\beta$ of 0.79 fits the pattern well in all generations.

The rare surnames in this sample are all associated with wealth. We can form from the Oxbridge records another larger rare surname group which consists just of rare surnames that show up as entrants to Oxbridge 1800-29. Table 1 also shows the relative representation of these surnames at Oxbridge to 2010. Here there is a large decline between 1800-29 and 1830-59. But to measure the true implied $\beta$ it is necessary to start with the generation 1830-59, where the elite surnames were selected based on their occurrence earlier, and so the data is not contaminated by positive errors. As can be seen this group also remains an elite even to 1980-2010. We can also calculate the implied $\beta$ for the regression to the mean of this group 1830-59 to 1980-2010, assuming as before that the initial variance in status was the same as for the population. It is 0.765, as is shown in figure 5. As before there is no sign of any speeding up of the process in the most recent generations.

Suppose we instead assume that the status variance of the rare surname group observed at Oxbridge in 1800-29 is instead 0 in 1830-59. How would that change the estimated $\beta$ to best fit the observed pattern of relative representation? Figure 6 shows the fitted path in this case that again minimizes the sum of squared deviations. Here the fit is less good. Such an assumption about initial variance implied a much

---

4 Judged by minimizing the sum of squared deviations (in logs).
more rapid initial decline in relative representation, which is not consistent with the data. However, the implied $\beta$ that best fits the observed pattern changes hardly at all. It goes from 0.765 to 0.770. So if we use the pattern of relative representation over many generations to estimate the implied $\beta$, even though we have to make an assumption about the initial variance in status of the elite, that will have little effect on the estimated value of $\beta$. In the results below I have thus assumed that the variance in status of elite or underclass groups always equals that of the population.
Results

Studies of surname distributions among elites and underclasses as outlined above for a variety of countries and time periods suggest the following hypotheses about social mobility rates.

1. The simple equation

\[ y_{t+1} = \beta y_t + u_t \]

where \( y_t \) is a measure of status, and \( u_t \) a random component, describes all social mobility. Elites and underclasses all tend to mediocrity at a constant rate. And the rate of movement is constant across generations.

2. \( \beta \) is much higher than conventionally estimated, between 0.7 and 0.8. Social mobility is extremely slow. Complete regression to the mean typically takes 10-16 generations, 300-500 years.

3. \( \beta \) is constant across societies and social systems.

4. \( \beta \) is constant across measures of status – wealth, education, occupation – and across the entire distribution of status, being the same for the upper tail as for the lower tale.

5. Since \( \beta^2 = 0.5-0.6 \) the majority of social status is determined at conception.

6. We observe persistent elites and underclasses only in two cases. The first is an isolated elite with marital endogamy (as with the Copts in Egypt). The second is where an elite or an underclass is maintained by selective retention of members with the elite or underclass characteristics, and recruitment of outsiders with the characteristic.

7. Assortative mating is what makes \( \beta \) so high. Mating has become more assortative in the modern world, so mobility rates may decline further (Herrnstein-Murray claim).

8. Social status is likely mainly of genetic origin.

Table 2 summarizes the results so far of the various \( \beta \) estimates by type of status, and by country and era. This summarizes the evidence for hypotheses 2-5 above. Below I go through each case to illustrate the various hypotheses. The “?” indicate cases where I hope to be able to derive estimates also.
Table 2: Estimates of $\beta$ from Surnames

<table>
<thead>
<tr>
<th>Country</th>
<th>Period</th>
<th>Wealth</th>
<th>Education</th>
<th>Occupations</th>
</tr>
</thead>
<tbody>
<tr>
<td>England</td>
<td>1800-2011</td>
<td>0.72</td>
<td>0.77</td>
<td>0.69</td>
</tr>
<tr>
<td>England</td>
<td>1300-1550</td>
<td>0.65</td>
<td>0.77</td>
<td>-</td>
</tr>
<tr>
<td>USA</td>
<td>1940-2010</td>
<td>-</td>
<td>-</td>
<td>0.74</td>
</tr>
<tr>
<td>Sweden</td>
<td>1650-2010</td>
<td>-</td>
<td>?</td>
<td>0.76</td>
</tr>
<tr>
<td>Bengal</td>
<td>1900-2010</td>
<td>-</td>
<td>-</td>
<td>0.80</td>
</tr>
<tr>
<td>Japan</td>
<td>1940-2011</td>
<td>-</td>
<td>0.84</td>
<td>0.82</td>
</tr>
<tr>
<td>Chile</td>
<td>1920-1990</td>
<td>-</td>
<td>?</td>
<td>0.74</td>
</tr>
<tr>
<td>China</td>
<td>1905-2011</td>
<td>-</td>
<td>0.75</td>
<td>?</td>
</tr>
</tbody>
</table>

Rare Surnames, England, 1780-2011

Groups of rare surnames, those held by 40 or less people in 1881, were identified as Very Rich, Rich, Average and Poor based on the average value of the probates of those dying with the surnames aged 21 and over in 1858-1887.\(^5\) We then measure the average wealth of these surnames for each of four subsequent death generations, 1888-1917, 1918-1952, 1953-1989, 1990-2024. Probate records give an indication of the wealth at death of everyone in England and Wales by name 1858 and later.\(^6\) The generations were allocated on the assumption that the average child was born at age 30 of the parent. The average child would thus die 30 years later, plus any gain in average years lived by adults of that generation.

These surnames are so rare that those of the rich are not generally perceived as having any special social status. Thus the first very rich names, in alphabetical order, were Ahmuty, Allecock, Angerstein, Appold, Auriol, Bailward, Bazalgette. The poor names, in order, begin Aller, Almand, Angler, Anglim, Annings, Austell, Backlake. Table 3 summarizes the data for each death period for the rich and poor. Figure 7 shows the average log probate value, by generation of death, for the initial surname groups and their descendants.

\(^5\) Clark and Cummins, 2012, describes this sample and the results in detail.
\(^6\) Those not probated typically have wealth at death close to 0.
Table 3: Summary of the Sample

<table>
<thead>
<tr>
<th>Period</th>
<th>Surnames</th>
<th>Probates</th>
<th>Deaths</th>
<th>Deaths 21+</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>VERY RICH/RICH</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1858-87</td>
<td>181</td>
<td>1,142</td>
<td>2,263</td>
<td>1,767*</td>
</tr>
<tr>
<td>1888-1917</td>
<td>172</td>
<td>1,072</td>
<td>1,987</td>
<td>1,792</td>
</tr>
<tr>
<td>1918-1952</td>
<td>168</td>
<td>1,582</td>
<td>2,478</td>
<td>2,383</td>
</tr>
<tr>
<td>1953-89</td>
<td>156</td>
<td>1,310</td>
<td>2,008</td>
<td>1,983</td>
</tr>
<tr>
<td>1990-2011</td>
<td>143</td>
<td>564</td>
<td>989</td>
<td>980</td>
</tr>
<tr>
<td><strong>MIDDLING/POOR</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1858-87</td>
<td>273</td>
<td>107</td>
<td>3,300</td>
<td>1,798*</td>
</tr>
<tr>
<td>1888-1917</td>
<td>255</td>
<td>275</td>
<td>3,106</td>
<td>1,889</td>
</tr>
<tr>
<td>1918-1952</td>
<td>242</td>
<td>638</td>
<td>3,085</td>
<td>2,610</td>
</tr>
<tr>
<td>1953-89</td>
<td>246</td>
<td>1,305</td>
<td>3,776</td>
<td>3,654</td>
</tr>
<tr>
<td>1990-2011</td>
<td>214</td>
<td>836</td>
<td>2,165</td>
<td>2,135</td>
</tr>
</tbody>
</table>

Note: * Where age was unknown 1858-65, the fraction above 21 was estimated from the 1866-87 ratio of deaths 21+ to all deaths.
We can estimate the \( \beta \) s, for wealth, in several different ways. If we define \( \bar{y}_{Rt} \) and \( \bar{y}_{Pt} \) as the average of ln normalized wealth for generation \( t \) for the rich and middling/poor surname groups, then the \( \beta \) linking this generation with the \( n \)th future generation can be measured simply as

\[
\bar{y}_{Rt+n} - \bar{y}_{Pt+n} = \beta(\bar{y}_{Rt} - \bar{y}_{Pt})
\]  

(5)

This measure will be, as described above, in expectation the same as the traditional intergenerational \( \beta \) estimates.

This estimation has an advantage described above that after the first generation, when rich and poor samples were chosen partly based on wealth, there is no tendency for the \( \beta \) estimate to be attenuated by measurement error in wealth, since the average measurement error for both rich and poor groups will be zero. Table 4 shows the implied \( \beta \)s, along with bootstrapped standard errors. Table 4 suggests two things. One is that the average \( \beta \) values between generations are much higher than are conventionally estimated. The average \( \beta \) value across 4 generations is 0.72. These values are so high that there is still a significant connection between wealth 4 generations after the first.
Table 4: $\beta$ Values Between Death Generations, England

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1858-1887</td>
<td>0.71 (.03)</td>
<td>0.62 (.02)</td>
<td>0.42 (.02)</td>
<td>0.26 (.03)</td>
</tr>
<tr>
<td>1888-1917</td>
<td>0.86 (.03)</td>
<td>0.59 (.03)</td>
<td>0.36 (.04)</td>
<td></td>
</tr>
<tr>
<td>1918-1952</td>
<td></td>
<td>0.68 (.03)</td>
<td>0.41 (.05)</td>
<td></td>
</tr>
<tr>
<td>1953-1987</td>
<td></td>
<td></td>
<td>0.61 (.07)</td>
<td></td>
</tr>
</tbody>
</table>

Note: Standard errors in parentheses.

The second suggestion of table 4, however, is that the $\beta$ may have fallen for the last generation, those dying 1999-2011. However, there is other evidence that suggests little increase in the rate of mobility in recent generations, and clear evidence that complete equality between the original rich and poor in wealth at death will not be accomplished before 2100. One element of this evidence is the continued overrepresentation of these surnames at Oxford and Cambridge as described above, which implied that even 60-70 years from now, in 2070-80, at the death of the current Oxbridge entrants, they will have higher wealth than those with non-elite surnames. Another element, outlined in Clark and Cummins (2012), is that if we organize the surname groups instead into birth cohorts, which gives us 6 generations of surnames born 1780-1809, 1810-39, 1840-69, 1870-99, 1900-29, and 1930-59, there is little sign of any decline in $\beta$ over time, and in particular for the last generation born 1930-59. Table 5 shows these estimates of $\beta$ from birth cohorts.
Table 5: $\beta$ values between birth generations, 1780-1809 to 1930-1959, England

<table>
<thead>
<tr>
<th></th>
<th>1810-39</th>
<th>1840-69</th>
<th>1870-99</th>
<th>1900-29</th>
<th>1930-59</th>
</tr>
</thead>
<tbody>
<tr>
<td>1780-1809</td>
<td>0.72</td>
<td>0.54</td>
<td>0.41</td>
<td>0.23</td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.04)</td>
</tr>
<tr>
<td>1810-39</td>
<td>0.75</td>
<td>0.57</td>
<td>0.32</td>
<td>0.22</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td></td>
<td>(0.06)</td>
</tr>
<tr>
<td>1840-69</td>
<td></td>
<td>0.76</td>
<td>0.41</td>
<td>0.29</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.03)</td>
<td>(0.03)</td>
<td></td>
<td>(0.07)</td>
</tr>
<tr>
<td>1870-99</td>
<td></td>
<td></td>
<td>0.56</td>
<td>0.39</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.04)</td>
<td></td>
<td>(0.10)</td>
</tr>
<tr>
<td>1900-29</td>
<td></td>
<td></td>
<td></td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.18)</td>
</tr>
</tbody>
</table>

Notes: $\beta$ values corrected to a 30 year generation gap. Standard errors were bootstrapped.

Why is Regression to the Mean so Slow in England, 1780-2011?

The $\beta$s we find here for wealth and education are high compared to the conventional estimates for the UK. It is this which allows for a significant connection between the wealth and educational attainment of people and their descendants 5-7 generations later. Table 6 shows a summary of recent estimates for $\beta$ for the UK.

Because of the design of the surname sample it oversamples the rich, particularly in the early years. Could it be that regression to the mean is slower for the very rich than for the population as a whole? We can rule out this possibility for wealth, however. Our data suggests the rate of regression to the mean is similar for the very rich, the rich and the poor. Table 7 thus shows separately for the very rich, the rich, and the poor the implied rate of regression to the mean in wealth between
Table 6: Modern Intergenerational Elasticities for the UK

<table>
<thead>
<tr>
<th>Measure</th>
<th>β</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Earnings</td>
<td>.22-.69</td>
<td>Dearden et al. (1997), Nicoletti and Ermisch (2008)</td>
</tr>
<tr>
<td>Wealth</td>
<td>.48-.59</td>
<td>Harbury and Hitchens (1979)</td>
</tr>
<tr>
<td>Education</td>
<td>.43-.71</td>
<td>Dearden et al. (1997), Hertz (2007)</td>
</tr>
<tr>
<td>Occupation</td>
<td>.08-.30</td>
<td>Francesconi and Nicoletti (2005), Ermisch et al. (2006)</td>
</tr>
</tbody>
</table>

Notes: Education refers to years of education, occupation to an index of occupational prestige (the Hope-Goldthorpe score).

Table 7: Average β versus “Brown” by Initial Wealth, England

<table>
<thead>
<tr>
<th></th>
<th>Gen 0 to Gen 1</th>
<th>Gen 1 to Gen 2</th>
<th>Gen 2 to Gen 3</th>
<th>Gen 3 to Gen 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Richest</td>
<td>0.72</td>
<td>0.68</td>
<td>0.79</td>
<td>0.66</td>
</tr>
<tr>
<td>Rich</td>
<td>0.78</td>
<td>0.87</td>
<td>0.79</td>
<td>0.62</td>
</tr>
<tr>
<td>Poor</td>
<td>0.73</td>
<td>0.40</td>
<td>1.70</td>
<td>0.84</td>
</tr>
</tbody>
</table>

the generation dying 1858-1887, and that dying 1999-2011, where we take as the base group the surname Brown(e), and estimate β from

\[
\bar{y}_{Rt+1} - \bar{y}_{Bt+1} = \beta (\bar{y}_{Rt} - \bar{y}_{Bt}) \quad (6)
\]

The average estimated β is 0.72 for the richest, 0.78 for the rich, and 0.73 for the poorest. There is no sign that slow regression to the mean is just a phenomenon of the very rich. Instead the β’s are remarkably similar across groups. Because, however, the poor were much closer in average wealth to the brown(e) surname, the
estimates of $\beta$ for this group are much less precise, and jump around from period to period. We also see in the Oxbridge data that as the wealth of the rich group becomes closer to the average in later generations, there is no sign of a speeding up of the decline of this group of surnames as an educational elite.

Using the census, birth, death and marriage registers for England 1837-2011 we are able to link 1,342 adult children to fathers, much of the rare surname sample. This was done by cross referencing the probate and death records with the census enumerator forms from 1841 to 1911. After 1911, all marriage index records listed the maiden name of the bride. In addition, all birth index records contained the maiden name of the child’s mother. It was thus possible to link children to marriages. Following this, marriages were linked to death and probate records. All ambiguous matches, where there was more than one potential match, were dropped. Using this data we can estimate directly the $\beta$ in

$$y_{ij,t+1} = a + \beta y_{it} + u_{ij,t+1} \quad (7)$$

where $y$ is log wealth at death, for individual families. Because the daughters observed are just those who were single at death (so retaining the family surname) we estimate

$$\ln(\text{WEALTH CHILD}) = a + \beta \ln(\text{WEALTH FATHER}) + \text{DFEM} + e$$

where DFEM is an indicator variable, 1 for a daughter. Table 8 shows these estimates for children dying in each of our death generations, compared to our estimates of $\beta$ from surname cohorts. The $\beta$ estimated from surnames is consistently higher. Table 9 shows the equivalent estimates when we organize children into birth cohorts. Again the $\beta$ estimated from direct family links is lower in each period.

---

7 The coefficient on the indicator variable for daughters is always negative.
Table 8: $\beta$ from Surnames and Families, by death generation, England

<table>
<thead>
<tr>
<th>Child Death Period</th>
<th>Surname Types $\beta$</th>
<th>Individual Surnames $\beta$</th>
<th>Linked Children Number</th>
<th>Individual Families $\beta $</th>
</tr>
</thead>
<tbody>
<tr>
<td>1888-1917</td>
<td>0.71</td>
<td>0.66</td>
<td>202</td>
<td>0.59</td>
</tr>
<tr>
<td>1918-1952</td>
<td>0.86</td>
<td>0.71</td>
<td>466</td>
<td>0.65</td>
</tr>
<tr>
<td>1953-1987</td>
<td>0.68</td>
<td>0.60</td>
<td>389</td>
<td>0.51</td>
</tr>
<tr>
<td>1988-2011</td>
<td>0.61</td>
<td>0.53</td>
<td>239</td>
<td>0.29</td>
</tr>
<tr>
<td>Average</td>
<td>0.72</td>
<td>0.62</td>
<td>-</td>
<td>0.51</td>
</tr>
</tbody>
</table>

Table 9: $\beta$ from Surnames and Families, by birth generation, England

<table>
<thead>
<tr>
<th>Child Birth Period</th>
<th>Surname Types $\beta$</th>
<th>Individual Surnames $\beta$</th>
<th>Linked Children Number</th>
<th>Individual Families $\beta $</th>
</tr>
</thead>
<tbody>
<tr>
<td>1810-39</td>
<td>0.72</td>
<td>0.63</td>
<td>72</td>
<td>0.13</td>
</tr>
<tr>
<td>1840-69</td>
<td>0.75</td>
<td>0.57</td>
<td>409</td>
<td>0.57</td>
</tr>
<tr>
<td>1870-99</td>
<td>0.76</td>
<td>0.71</td>
<td>440</td>
<td>0.57</td>
</tr>
<tr>
<td>1900-29</td>
<td>0.56</td>
<td>0.48</td>
<td>242</td>
<td>0.36</td>
</tr>
<tr>
<td>1930-59</td>
<td>0.69</td>
<td>0.31</td>
<td>178</td>
<td>0.26</td>
</tr>
<tr>
<td>Average</td>
<td>0.70</td>
<td>0.54</td>
<td>-</td>
<td>0.44*</td>
</tr>
</tbody>
</table>

Notes: *birth generations 1840-69 to 1930-59
The estimates in tables 8 and 9 suggest that the reason for the estimated slow regression to the mean of our surname groups is not because of any unusual persistence of wealth in England by conventional standards. The $\beta$s estimated for the family linkages for recent years are at the low end of the range reported for the modern UK in table 6.

Conventional estimates of the intergenerational elasticity of wealth, earnings and education, once corrected for measurement errors, correctly answer the question about what the inheritance of any of these aspects of underlying status is in one generation. But they will underestimate the intergenerational elasticity of any of these aspects of status across subsequent generations, and hence of what long run social mobility is on this particular measure. They will also underestimate what the intergenerational elasticity of a broader measure of socio-economic status, which averaged wealth, earnings, education, health, and other aspects of status is. These measures thus systematically overestimate social mobility.\(^8\)

By switching to surname cohorts, we are instead measuring an intergenerational elasticity closer to this long run elasticity, and to the elasticity of broader measures of status. The estimates of $b$ in columns 3 of tables 14 and 15, using surname families, are higher than those in column 5 which use actual families in part because by averaging wealth across larger groups of people with the same surname we are reducing the random element in wealth estimates in the first generation.

This difference in the two measures of intergenerational elasticity shows up in the last column of table 9, showing Oxbridge attendance among rare surnames. If we identify a rare surname elite by entry to Oxbridge 1800-29 of someone with this surname, then the implied intergenerational elasticity for entry to this elite between this first generation and the next is 0.61. But for all subsequent generations the implied elasticity drops to 0.76.

---

\(^8\) Confirming this Jason Long in a study of occupational mobility in England was able to link sons, fathers and grandfathers in 1851, 1881, and 1901 (Long, 2011). Even controlling for the occupation of fathers, the occupation of grandfathers was predictive of the occupation of sons. There was more persistence of occupational status than the one generation elasticity would suggest (personal communication from author).
England, 1086-1780

I can also use surnames to estimate social mobility in medieval England. Surnames in England spread downwards from the upper classes to the lower in the period 1066-1300. By 1300 most (but not all) of the English had surnames, but for most people in 1300 their surnames were of recent vintage, no more than 2-3 generations from the original bearer. In England in 1200-1300 surnames differed significantly in average social status. There were aristocratic names, typically derived from the location of the main residence of the family, such as (de) Berkeley and (de) Merton, or even of the residence of the family before the Norman Conquest such as Baskerville, and Neville. There were names derived from occupations of high socio-economic status, such as Chamberlain, or Stewart, or Clark. And there were surnames denoting occupations in the middle of the social hierarchy, such as Smith, Cook, Baker, and Wright.

Reflecting these differences in social status surnames appeared among the students and faculty of Oxford and Cambridge universities at very different rates relative to their share in the population. Thus despite the fact that from 1540 on artisan surnames constituted 7% of the surname stock of England, before 1270 there were no such surnames among those attending Oxbridge. Figure 10 shows the relative representation of artisan surnames oat Oxbridge by generation 1230-1950. In the 200 years 1270-1470 artisan surnames rose from no representation to being fully represented in the universities. There was a complete absorption of the descendants of the artisans into the educational elite by 1470, before the end of the middle ages, and before the reform of the Catholic Church in the English Reformation.

Figure 10 gives the appearance of an astonishingly rapid rise of the descendants of the artisans, the sons of the Smiths, in the late middle ages. But the implied rate of persistence, the $\beta$ for medieval England, is quite high. To estimate this I assume initially that the artisans lay at the mean of socio-economic status in 1290-1319, that they all had the same social status so that the variance across individuals in these occupations was 0, and that Oxbridge represented the top 1% of the social hierarchy. With these assumptions, as figure 11 shows, the best estimate of the $\beta$ for Medieval England is 0.75.

This assumes, however, that the “Smiths” and other artisans coming of age in 1290-1319 were all the first generation holders of the surname. A more conservative assumption is that the typical Smith acquires their surname in the previous generation, 1260-89. In that case the best fit for the pattern of representation at Oxbridge is a $\beta$ of 0.8. If I was to assume that the artisans lay a bit lower than the mean in social status initially then the implied $\beta$ would be lower. If I assumed that in the beginning there was some distribution of status among artisans then the implied $\beta$ would be higher. But overall a $\beta$ of 0.75-0.8 seems to fit well the medieval data. This implies that rates of social mobility in medieval England were as high as they are now in England.
Figure 10: Artisan and Upper Occupation Surnames at Oxbridge, 1170-1950

Figure 11: Fitted Path of Regression to the Mean, Artisans, 1230-1590
Figure 10 also shows the relative representation of “elite” occupations at Oxbridge in medieval England, where such occupations were those of high manorial officials such as Chamberlains, government officials and attorneys (Clarks), or high commercial occupations such as clothier, draper, mercer, and merchant. Though the data is noisier such occupational surnames achieve a relative representation at Oxbridge of 1 even sooner than the surnames of lower class occupations. By 1380-1409 they are fully represented at the universities.

The rise of the artisan class was accompanied by the decline of the traditional landed classes. For 1230-99 I identify elite surnames through the records of those dying in possession of lands held of the king, through a royal set of records enquiring into such deaths, the Inquisitiones Post Mortem. Figure 12 shows the relative representation of a sample of surnames occurring in the Inquisitiones post Mortem 1230-99 at Oxbridge 1200-1600. The figure shows that in the 13th century these times were four times as frequent at Oxbridge as in the general population, but by the 16th century they were very little overrepresented at the universities. The figure also shows the best fitted $\beta$ for this pattern assuming that the variance in status of this elite was as great as in the general population, and that the peak of status was achieved in 1230-59. The best fitting $\beta$ is 0.78, which is remarkable congruent with the $\beta$ for the medieval artisans, as well as the $\beta$ for elites at Oxford and Cambridge 1800-2011.
The Rise of Elites

We see above that in both modern and medieval England one simple equation

\[ y_{t+1} = \beta y_t + u_t \]

with a constant \( \beta \) seems to describe social mobility at all levels of the society. There seems to be a simple physics of social mobility. One implication of this equation is that social mobility should be the same forwards and backwards. Surname information should dissipate at the same rate forwards and backwards. The move of those at the extremes of the distribution – extremes of wealth or poverty, education and ignorance – towards the center should be symmetrical with their earlier move from the center to the extremes. Any group at the extreme should not only regress to the mean in future generations, it should also seem to regress to the mean in the same fashion if we go back through earlier generations.

Using the database of Oxford and Cambridge students we can test this proposition for England. I start with the period 1800-29, and look at all students at Oxford and Cambridge where 40 or fewer persons were recorded with that name in the 1881 census. People with these rare surnames thus represent on average an educational elite when observed in 1800-29. Figure 13 shows their relative representation at Oxford and Cambridge for each 30 year generation 1830-59, …, 1980-2009, and also 2010-14. There is the predicted period by period decline in relative representation for the 7 generations following. But interesting these names show a near symmetrical rise in the 7 generations before 1800. Complete symmetry is indicated by the dotted line in figure 14 (done in logs). So clearly the path does not achieve this. But the identification of the stocks of people with the rare names in the years before 1800, necessary to estimate the relative representation, is tentative at present.\(^9\) The figure is nonetheless powerful testament to the simple equation used in this study being close to the true. Elites arise step by step from mediocrity, have their generation in the sun, and then fade back to mediocrity. Though as figure 13 shows this is at least a 10 generation long process in either direction.

\(^9\) It is based on the frequency of surnames in marriages in England 1540-1800, using Boyds Marriage index. But in earlier years spelling is highly variable, and may be more accurate in the college records than in the marriage registers.
Figure 13: Relative Representation of an Elite in 1800-29 over 15 Generations at Oxbridge

Figure 14: Figure 13 in Logs
Sweden, 1600-2011: Surprising Rigidities

Modern Sweden is known to be a place of high degrees of intergenerational income mobility. Thus Black and Devereux (2010) report in their survey of mobility across countries that the b for income in Sweden for men of 0.26, and for women of 0.19. This would imply that after two generations someone whose income was double the average would have grandchildren who had incomes only 1-2% above the average. Suppose these estimates are indicative of general social mobility in Sweden over the last three generations in which the social democrats were important actors in the political system, 1917-2011. Then there should be little connection between the social status of the current generation, and the status of Swedes before 1914.

Such enhanced mobility in a country like Sweden would also suggest that institutional arrangements – the support for public education, for example, or the progressive taxation of wealth - play a vital role in determining rates of social mobility. The implication is that the lower rates of social mobility observed in countries such as England or the USA represent a social failure. The life chances of the descendants of high and low status ancestors can be equalized at low social cost. Sweden is, after all, one of the richest economies in the world.

Here I show, however, that at least in the case of Sweden the true intergenerational mobility of status was likely no greater than that of England over the past 200 years. Whatever the short run mobility of income, there is considerable persistence of status – measured through wealth, education and occupation - over many generations in Sweden.

To look at long run mobility in Sweden I employ three types of surnames: surnames of nobles, Latinized surnames, and patronyms. In the sixteenth to eighteenth century, when most Swedes did not have inherited surnames, the educated class – clerics, academics, and some merchants - adopted Latinized surnames (typically ending “ius” or “aeus”), which became characteristic of them as a class.

Noble Surnames

Sweden has a formal guild of noble families, the Riddarhuset (House of Nobility). Though noble families existed since medieval times, the modern Riddarhuset was created in 1626. During the Diet of the Four Estates, 1668-1865, the Riddarhuset functioned as one of the four governing estates of the kingdom (analogous to the House of Lords in England).10

---

10 The organization has an elaborate building in Stockholm, also called the Riddarhuset, erected in the seventeenth century, which functioned 1688-1865 as one of the houses of the Parliament.
Since 2003 the Riddarhuset has been a purely private institution, which maintains the records of the Swedish noble families, and lobbies on their behalf.

The families enrolled in the Riddarhuset come in three ranks: counts, barons, and “untitled” nobility. Each family has a number corresponding to their order of enrollment. In total 2,330 families have been enrolled, though only about 700 have living representatives. Though the Swedish King could create nobles up until 1975, these ennobled had to independently seek enrollment in the Riddarhuset. The last person ennobled by the King was in 1902.

The great period of expansion of the Swedish mobility was in 1626-1658, when Sweden enjoyed a period of conquest that brought its territories to their maximum extent of 1658-1721. At this point Sweden had possession also of Finland, Estonia, and some north German states. In this period the crown rewarded many military commanders with ennoblement. This is reflected in the foreign (particularly German) names of many of the nobility, who served the crown as military commanders.

From 1680 the nobility gradually lost its privileges, starting with the reclamation by the crown in 1680 of much of the land granted to nobles in previous years. By 1866 the nobles had no privileges on any economic significance.

When families were enrolled in the Riddarhuset they typically adopted a new surname if they were Swedish, embodying status elements such as “Gyllen” (gold), “Silfver” (silver), “Adler” (eagle), “Leijon” (lion). Many of the noble surnames in Sweden, however, are German in origin, reflecting the importance of German military commanders in the service of the Swedish crown in the seventeenth century.

One important privilege that the nobility obtained in the Names Adoption Act of 1901 was a ban on anyone else adopting their surnames.11 Thus apart from foreign imports, and name changing before 1901, the surnames of the enrolled nobles in the Riddarhuset identify uniquely the lineage of these noble families. The modern generation with these surnames are all descended from a privileged class dating typically from 1721 and before.

Latinized Surnames

A second category of surnames in Sweden which reveals that the holder had ancestors of higher social status are those that have been Latinized. Such names would have been typically adopted as family surnames before 1800 by educated people. This is reflected in the names of a number of famous Swedish scientists of the seventeenth and eighteenth

---

11 There had been concern that disreputable people had been adopting noble surnames.
centuries: Carolus Linnaeus (1707-78), Anders Celsius (1701-44), Jöns Jakob Berzelius (1779-1848), Olaus Rudbeckius (1630-1702), Olaus Rudbeckius (Junior) (1660-1740).

Indeed, taking a sample of such surnames held now by 200 people or more – Afzelius, Ahlenius, Alenius, Arrhenius, Axtelius, Bergius, Bruzelius, Forselius, Helenius, Landelius, Montelius, Sandelius, Stenius – the date when the name first appears in just a sample of Swedish parish records of baptisms, burials and marriages was: 1669, 1745, 1646, 1630, 1749, 1591, 1747, 1626, 1666, 1743, 1668, 1682, 1724.12

Patronymics

The lowest class of surnames in Sweden are those ending in “son”. This is because these were the surnames adopted in the late nineteenth century by the lower classes in Sweden. And they are still of a distinctly lower status than surnames which embody topographic or natural elements such as those containing “lund”, “berg”, “qvist”.

Surname Changing

Are modern Swedes with Latinized surnames mostly the descendants of the clerical and professional classes of 1550-1800? A potential problem here is that for many Swedes before the 1860s a system of patronyms, such as Andersson, changing by generation with the first name of the father, served as surnames. In the late nineteenth century these transient patronyms were translated into fixed surnames. In this process did many people adopt Latinized surnames, as opposed to ossifying their current transient patronym?

We can examine this question by looking at frequency changes for names between the censuses of 1880 and 1900. Between these two dates the numbers of Anderssons reported in the census rose by 23%. Over the same interval the numbers of those with the above reported 13 Latinized surnames rose modestly more, by 27%. This could easily be just differences in birth rates between these groups, or differences in emigration rates. So there is no sign in the late nineteenth century of any large scale switch to Latinized surnames.

After 1901 surnames in Sweden became much more rigidly attached to people with the arrival of the formal system of surname registration in the 1901 Names Adoption Act. This

---

12 The sample names were checked against the International Genealogical Index of the Church of Jesus Christ of Latter-Day Saints, searching under the first names Johan, Anders, Karl, Erik and Nils (http://www.familysearch.org/eng/search/frameset_search.asp). These records give only a limited sample of parish records in Sweden before 1800.
abolished the system of generation specific patronyms, and required each family to have an unchanging surname. Only in 1982 did it become possible for people to change their surname. But the 1982 law that permits this establishes restrictive criteria for registering a new surname. People are not allowed to change their surname to one held by existing families, with two exceptions. They are allowed to reclaim surnames that have been used by their ancestors for at least two generations within the last 100 years. And they are allowed to produce spelling variants of their surname (“Andersson” can be changed to “Anderson” for example).\[13\]

However, among the new surnames the current law allows people to adopt are those with the Latinized ending “ius”. Thus some of the “ius” ending names have potentially been created since 1982. But since the law also demands uniqueness in the formation of new surnames, the numbers of people holding each such neologistic surnames will be less than 10 in almost all cases. I can thus control for the effect of such recent creations by also confining the Latinized surnames to those held by 10 or more people in 2010.

### Social Mobility Rates

Figure 15 shows the frequency per 1,000 of each surname type in the population of physicians in Sweden registered in 2010. The surnames of the high nobles, the counts and barons, largely created by 1700 are still the most frequent of all Swedish surnames among physicians in 2010. Indeed they are 6 times as frequent per 1,000 as those surnames ending in “son”. But the “untitled” nobility, the Latinized surnames, and surnames containing “lund” are all more heavily represented among doctors than the “son” names. Since there are many foreign physicians in Sweden, estimating the average frequency among doctors of Swedish origin surnames is tricky. But the “lund” surnames seem to come close, and these are taken as the reference. Figure 15 implies that there has been very slow occupational mobility in Sweden over the years 1700-2010. Ten generations after the creation of Latinized and Noble surnames they are still strongly overrepresented among the modern elite. The mobility displayed here cannot be any higher than that revealed in England through the Oxbridge sample, and might even be lower.

Since the medical registers give the date of admission of each person, it is possible for Sweden to construct decadal estimates of new doctors by surname type 1890-2010. Taking doctors as representing the top 1% of the occupational ranking, from this I can calculate the implied $\beta$ by generation for the last three generations of Swedes entering medical training. Table 10 shows this result for the Latinized surnames, and the patronyms, taking “lund” containing surnames as the reference.

---

\[13\] New surnames also have to be unique. Surnames also cannot end in “son.”
Figure 15: Physician Frequency by Surname Type, Sweden, 2010

Table 10: Implied $\beta$s for Sweden from Physician Surnames, 1920-2009

<table>
<thead>
<tr>
<th>Year</th>
<th>Implied $\beta$</th>
<th>Implied $\beta$</th>
<th>Implied $\beta$</th>
<th>Implied $\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1920-49</td>
<td>5.17</td>
<td>0.16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1950-79</td>
<td>3.44</td>
<td>.72</td>
<td>0.29</td>
<td>0.70</td>
</tr>
<tr>
<td>1980-2009</td>
<td>2.31</td>
<td>.65</td>
<td>0.38</td>
<td>0.79</td>
</tr>
</tbody>
</table>
Despite Sweden’s reputation for greater social mobility than in societies such as the UK or USA, the implied βs, calculated from the relative representation of the surnames using what seems like an entirely reasonable assumption that in each period the variance of occupational status for each surname type is the same as that of the population, are high. For the generation entering medicine 1950-79 they average 0.71, while for the generation entering 1980-2009 they average 0.72. These are very similar to the βs estimated above for modern and medieval England. If we look at the pattern of relative representation by decade for the elite groups then the impression is that social mobility has perhaps even slowed in the last few decades. Figure 16 shows this for the higher and lower aristocrat surnames, and the Latinized surnames. Between the 1970s and the 2000s for these groups the implied β would be 1 or greater. Modern Sweden may thus have lower social mobility rates than medieval England. However, the medical training in Sweden is rationed by the universities because of excess student demand. So this appearance of rigidity may owe not to a breakdown in the general law of regression to the mean, but to a temporary disruption caused by changes in administrative decisions about how to ration access to medical training.
The US as a nation of immigrants has a tremendous variety of surnames, and so is a promising place to investigate social mobility using surname distributions among elites and underclasses. There are four types of surnames that will be used here. Jewish surnames such as Cohen, Katz, Levin;\textsuperscript{14} surnames where at least 90 percent of the holders in 2000 identified themselves as Black, names such as Washington, Smalls, Merriweather, and Stepney; French surnames from Quebec such as Hebert, Cote, Gagnon; and (hopefully) rare English, Dutch and German surnames where at least one person with the surname attended an Ivy League college in the USA 1650-1850. Again the AMA directory of registered physicians is used to identify an elite occupational group.

Common surnames in the US still vary widely in average social status. Figure 17, for example, shows the relative representation of surname types in the AMA Directory of Physicians for 2010, where the vertical axis is on a log scale to allow for comparisons. Jewish surnames have a relative representation of 5, while Black surnames are at 0.3, and Native American surnames at 0.1.\textsuperscript{15} Since the AMA directory lists the date of registration of the current physician stock (including many retired and even deceased physicians), it shows cohorts of doctors by surname registering between the 1930s and 2000s. Thus it is possible to observe the implied social mobility of these surname groups over the last 60 years in the USA.

Table 11 shows the implied $\beta$s over the past two generations for the Jewish and Black surnames. The Jewish surnames were actually increasing their share of physicians relative to their share of the population between the generation first registering 1920-49 and that first registering 1950-79. However, the generation 1980-2010 shows a decline in relative representation, and an implied persistence $\beta$ of 0.82. However, if we go to the decadal level (where we can still estimate an implied $\beta$, which has to be cubed to give the generational $\beta$), the implied regression to the mean of the Jewish surnames in the US has been more rapid in recent decades. Figure 18 shows the relative representation of these surnames by decade from the 1970s, where the relative representation peaked, to the 2000s. Table 12 shows the calculated implied $\beta$s over these decades. For the Jewish surnames the average implied $\beta$ from the 1970s to the 2000s is 0.61.

For the Black surnames the generational estimates in table 11 suggest no regression to the mean between the generation entering medicine 1920-49 and 1950-79, and a $\beta$ of 0.70 between 1950-79 and 1980-2010. However, looking at the decadal level as in table 12, the

\textsuperscript{14} These Jewish surnames were identified as surnames of German origin now found at very low frequencies in Germany, but at high frequencies in New York State and Florida.

\textsuperscript{15} Black and Native American Surnames were identified from a report of the Census Bureau on the ethnicity of holders of all surnames of frequency 100 or greater in the 2000 census.
Figure 17: Relative Representation of Surname Types Among US Doctors, 2010

Table 11: Implied $\beta$s by Generation, Jewish and Black Surnames among Doctors

<table>
<thead>
<tr>
<th></th>
<th>Jewish</th>
<th>Implied $b$</th>
<th>Black</th>
<th>Implied $b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1920-49</td>
<td>4.53</td>
<td>-</td>
<td>0.13</td>
<td></td>
</tr>
<tr>
<td>1950-79</td>
<td>5.30</td>
<td>-</td>
<td>0.13</td>
<td>1.00</td>
</tr>
<tr>
<td>1980-2010</td>
<td>4.08</td>
<td>.82</td>
<td>0.25</td>
<td>0.70</td>
</tr>
</tbody>
</table>
Figure 18: Relative Representation, Jewish and Black Surnames among Physicians, USA

Table 12: Implied $\beta$s by Decade, USA, 1970s-2000s

<table>
<thead>
<tr>
<th>Decade</th>
<th>Jewish</th>
<th>Black</th>
</tr>
</thead>
<tbody>
<tr>
<td>1970-9</td>
<td>5.72</td>
<td>0.19</td>
</tr>
<tr>
<td>1980-9</td>
<td>4.96</td>
<td>0.22</td>
</tr>
<tr>
<td>1990-9</td>
<td>3.59</td>
<td>0.26</td>
</tr>
<tr>
<td>2000-9</td>
<td>3.30</td>
<td>0.28</td>
</tr>
<tr>
<td>Ave. b</td>
<td>0.61</td>
<td>0.78</td>
</tr>
</tbody>
</table>
implied upwards mobility of the Black surnames has been slower in recent decades, and implies over the last 4 decades a $\beta$ of only 0.78.

However, overall the physician data in the USA suggests that the upwards mobility of Black surnames in recent decades has been roughly the same as that of “son” surnames in Sweden. And the downwards mobility of Jewish surnames among physicians in the USA has been greater in recent decades than the downwards mobility of aristocratic and Latinized surnames in Sweden.

[TO BE ADDED – SECTION ON BENGAL, 1870-2011, JAPAN, 1870-2011, CHINA 1650-2011]
Data Sources

England

Medieval Elite:


Wealth:

*Index to the Prerogatory Court of Canterbury Wills*, 1384-1858. (UK National Archives, London).

*England and Wales, Index to Wills and Administrations, 1858-2011*. Principal Probate registry, London (available online 1861-1898, 1903-1942 at Ancestry.co.uk).

Births and Deaths:

*England and Wales, Register of Births, 1837-2005*. Available online at Ancestry.co.uk.

*England and Wales, Register of Deaths, 1837-2005*. Available online at Ancestry.co.uk.


University Attendance:


India


East India Company, List of Revenue Farmers, c. 1781, IOR/H/381


[http://ceowestbengal.nic.in/AcNoMatrix.asp?m_menu_id=40&m_district_id=11](http://ceowestbengal.nic.in/AcNoMatrix.asp?m_menu_id=40&m_district_id=11)

(downloaded and digitized by Lincoln Atkinson).

Kolkata Residential Phone Directory

[http://www.calcutta.bsnl.co.in/newdirectory/dqresidential.php](http://www.calcutta.bsnl.co.in/newdirectory/dqresidential.php)

Sweden

Attorneys: Members of the Swedish Bar Association

[http://www.advokatsamfundet.se/Advokatsamfundet-engelska/Find-a-lawyer/](http://www.advokatsamfundet.se/Advokatsamfundet-engelska/Find-a-lawyer/)

Surname frequencies, 2010: Statistics Sweden,  [http://www.scb.se/Pages/NameSearch](http://www.scb.se/Pages/NameSearch)


Swedish Censuses 1880, 1890, 1900: Riksarkivet,  [http://www.svar.ra.se/](http://www.svar.ra.se/)

Tax Assessments, 2008:

Taxerings- och förmögenhetskalendar för Stockholms kommun 2008.

Taxerings- och förmögenhetskalendar för Stockholms län Norra 2008.

Taxerings- och förmögenhetskalendar för Stockholms län Södra 2008.

[http://www.taxeringskalendern.se/](http://www.taxeringskalendern.se/)

University Theses:
Physicians

References


Jobling, Mark A. 2001. “In the name of the father: surnames and genetics.” *TRENDS in Genetics, Vol. 17, No. 6* (June): 353-357.


McCloskey, Deirdre. 2008. “You know, Ernest, the rich are different from you and me: A comment on Clark’s *Farewell to Alms.*” *European Review of Economic History, 12*(2) (August).


