Intergenerational Schooling Mobility
Current State of the Literature on Descriptive Estimates

Bradley J. Setzler

Economics 350
Department of Economics
University of Chicago
setzler@uchicago.edu

February 11, 2014
Four primary mechanisms have been proposed to explain the intergenerational mobility of schooling from parents to children:

- Genetics: parents pass unobservable ability to children, and both parents and children experience higher returns to schooling due to this ability;
- Parenting: educated parents are better at parenting, and these parents cultivate unobservable ability that raises returns to schooling for children;
- Credit constraints: less educated parents are also less wealthy and cannot afford schooling for children; and,
- Institutions: factors outside of the family like quality of local schools and discriminatory schooling laws determine schooling decisions.

Of course, some combination of these mechanisms is also possible.
To estimate the causal effects, the literature attempts to control for ability and estimate the causal marginal effect of parents’ education on children’s schooling using,

- differences in twin children’s years of schooling as a function of differences in parents’ years of schooling when twins are raised by different parents,
- regression of adopted children’s years of schooling on adopting parents’ years of schooling, or,
- compulsory schooling laws as instruments in IV regressions.

See Holmlund, Lindahl, and Plug (2011) for a recent survey of the causal effects methods used above. Not included in their survey are structural models of schooling decisions, which are covered in this class.
In contrast to these causal effects, the present survey is interested in descriptive parameters of intergenerational schooling mobility.

Before modeling the descriptive relationships in intergenerational educational mobility, recall the usual descriptive relationship estimated in studies of intergenerational earnings mobility. The intergenerational elasticity of earnings (IGE), originally derived by Becker and Tomes (1979) as the reduced form of an economic relationship and later modified to a relationship in log-earnings, is defined as $\beta_E$ in the regression,

$$\log E_i^c = \alpha_E + \beta_E \log E_i^p + X_i + \eta_i^E,$$

where $E$ is earnings, $X$ includes family characteristics, $c$ is children, $p$ is parents, and $i$ is the family index.
In contrast to the IGE, intergenerational schooling mobility has traditionally been estimated as the regression coefficient in the level regression,

$$S^c_i = \alpha + \beta S^p_i + X_i + \eta_i,$$

where $S$ is years of schooling (treated as a continuous variables), $c$ is child, $p$ is parent, and $i$ is family index. Sometimes the schooling years are normalized to unit variance so that $\beta$ is a conditional correlation.

Another parameter common in the literature is the correlation, often rank correlation, between parent and child years of schooling.
Alternatively, some define the parameter of interest as the marginal effect of parents’ educational attainment on probability of children’s attainment, that is, $\gamma$ in the regression moment,

$$F \left( \mathbb{E} \left[ A_i^c \mid A_i^p, X_i \right] \right) = \gamma G_i^p + X_i,$$

where $A \in \{0, 1\}$ indicates attainment of an education level and $F$ is typically the inverse standard normal CDF (probit regression) or the log odds-ratio (logistic regression).

Others extend this model to permit ordered categories corresponding to educational attainment levels, using ordered probit or logit. Still others ignore the natural order of educational attainment and estimate multinomial probit or logit.
To my knowledge, the first estimates of intergenerational elasticity of schooling were presented by Spady (1967) and Bowles (1972), who use the Occupational Changes in a Generation supplement to the 1962 Current Population Survey in the USA.

- Spady finds a rank correlation between son’s and father’s educational attainment levels ranging from 0.18 to 0.28, depending on father’s age.
- Bowles finds correlation between father’s and son’s educational attainment levels of 0.20 from a normalized regression.
Early USA Estimates: OCG

I use OCG 1973 to find the following correlations relating parents’ education to children’s education for whites:
Early USA Estimates: OCG

I find the following regression coefficients relating parents’ education to children’s education for OCG 1973 whites:

![Graph showing regression coefficients over time for different relationships between parents' and children's education.](image)
I use OCG 1973 to find the following correlations relating parents’ education to children’s education for blacks:
(Note: Sample sizes are only 38 for 1860 and 32 for 1930.)
I find the following regression coefficients relating parents’ education to children’s education for OCG 1973 blacks:
(Note: Sample sizes are only 38 for 1860 and 32 for 1930.)
Mare (1980) is the first to compare income correlations across generations using the Occupational Changes in a Generation supplement to the 1973 Current Population Survey in the USA. He is also the first to estimate continuation effects: he conditions on child’s education level, and estimates by logistic regression the conditional effect of both father’s and mother’s education level (measured in 16 grade categories, treated as a continuous variable) on child’s probability of continuing to the next level of schooling.

- For father’s education, he finds estimates ranging from 0.12 for finishing elementary school, to -0.01 for attending graduate school given college graduation;
- For mother’s education, he finds estimates ranging from 0.17 for finishing elementary school, to 0.04 for attending graduate school given college graduation;
- These effects are estimated to be strictly decreasing with each additional level of child’s education.
Behrman and Taubman (1986) regress the years of schooling for sons and daughters on years of education for their mothers and fathers in the NAS/NRC twins data, with little change in estimates depending on birth order:

- Coefficient on mother’s education for sons: 0.05-0.06;
- Coefficient on father’s education for sons: 0.16-0.17;
- Coefficient on mother’s education for daughters: 0.08-0.09;
- Coefficient on father’s education for daughters: 0.16-0.17.
Blake (1985) regresses child’s years of schooling on father’s years of education for three data sets, finding ranges of estimates depending on number of siblings:

- OCG 1962: 0.16-0.23;
- OCG 1973: 0.19-0.35;
- GSS 1972-83: 0.19-0.27.

Notice that the high end of these ranges is much larger than earlier estimates, even using the same OCG data. GSS is the General Social Survey.
Mulligan (1997) finds in regression using the SRC subset of the PSID data,

- Estimate of 0.32 for father’s education on sons; and,
- Estimate of 0.33 for father’s education on any children.
Couch and Dunn (1997) find that the following correlations in years of schooling using PSID:

- Regression coefficient (correlation) of mother’s education for sons: 0.37 (0.42);
- Regression coefficient (correlation) of father’s education for sons: 0.27 (0.42);
- Regression coefficient (correlation) of mother’s education for daughters: 0.32 (0.43);
- Regression coefficient (correlation) of father’s education for daughters: 0.26 (0.40).

These PSID estimates are much higher than the earlier estimates, especially the OCG estimates.
### Descriptive Estimates of Intergenerational Persistence in Years of Schooling in the USA

**Bradley J. Setzler**

<table>
<thead>
<tr>
<th>Source</th>
<th>Data</th>
<th>Method</th>
<th>Relationship</th>
<th>Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Spady (1967)</strong></td>
<td>OCG 1962</td>
<td>Condit. Rank Correlation</td>
<td>Father-Son</td>
<td>0.18/0.28</td>
</tr>
<tr>
<td><strong>Bowles (1972)</strong></td>
<td>OCG 1962</td>
<td>Normalized Regression</td>
<td>Father-Son</td>
<td>0.20</td>
</tr>
<tr>
<td><strong>Behrman &amp; Taubman (1985)</strong></td>
<td>NAS/NRC Twins 1981</td>
<td>Regression</td>
<td>Father-Son</td>
<td>0.17/0.19</td>
</tr>
<tr>
<td></td>
<td>OCG 1962</td>
<td>Normalized Regression</td>
<td>Father-Son</td>
<td>0.16/0.23</td>
</tr>
<tr>
<td><strong>Blake (1985)</strong></td>
<td>OCG 1962</td>
<td>Regression</td>
<td>Father-Son</td>
<td>0.16/0.17</td>
</tr>
<tr>
<td></td>
<td>OCG 1973</td>
<td>Regression</td>
<td>Father-Son</td>
<td>0.16/0.17</td>
</tr>
<tr>
<td><strong>Behrman &amp; Taubman (1986)</strong></td>
<td>NAS/NRC Twins 1981</td>
<td>Regression</td>
<td>Father-Son</td>
<td>0.16/0.17</td>
</tr>
<tr>
<td><strong>Hill &amp; Duncan (1987)</strong></td>
<td>PSID 1984</td>
<td>Regression</td>
<td>Father-Son</td>
<td>0.09/0.14</td>
</tr>
<tr>
<td><strong>Teachman (1987)</strong></td>
<td>NLSY 1979</td>
<td>Regression (Correlation)</td>
<td>Father-Son</td>
<td>0.18 (0.39)</td>
</tr>
<tr>
<td><strong>Couch &amp; Dunn (1997)</strong></td>
<td>PSID 1984/89</td>
<td>Normalized Regression (Correlation)</td>
<td>Father-Son</td>
<td>0.27 (0.42)</td>
</tr>
<tr>
<td><strong>Mulligan (1997)</strong></td>
<td>SRC (subset of PSID)</td>
<td>Regression</td>
<td>Father-Son</td>
<td>0.32</td>
</tr>
<tr>
<td><strong>Belzil &amp; Hansen (2003)</strong></td>
<td>NLSY 1979/90</td>
<td>Regression</td>
<td>Father-Son</td>
<td>0.21</td>
</tr>
<tr>
<td><strong>Farre, et al. (2012)</strong></td>
<td>NLSY 1979/90</td>
<td>Regression</td>
<td>Father-Son</td>
<td>0.21</td>
</tr>
</tbody>
</table>

Note: Ranges of results are designated by the forward slash and correspond to various controls included in regressions. See my survey, *The Intergenerational Persistence of Schooling: Current State of the Literature on Descriptive Estimates*, for citations, estimates for other countries, and estimates for measures of schooling attainment other than years of schooling (e.g., probability of college attendance).
Couch and Dunn (1997) appear to be the first to use comparable estimates between two countries, as they compare the PSID estimates above to these estimates from Germany’s SOEP data:

- Regression coefficient (correlation) of mother’s education for sons: \(0.11 (0.10)\);
- Regression coefficient (correlation) of father’s education for sons: \(0.20 (0.24)\);
- Regression coefficient (correlation) of mother’s education for daughters: \(0.48 (0.39)\);
- Regression coefficient (correlation) of father’s education for daughters: \(0.03 (0.02)\).

With the exception of the strong correlation between mother’s and daughter’s education in Germany, the USA is found to have much lower intergenerational mobility in years of schooling.
Dearden, Machin, and Reed (1997) regress years of education for sons and daughters on mother’s and father’s education for Britain’s NCSD data set, finding:

- Coefficient on mother’s education for sons: 0.25;
- Coefficient on father’s education for sons: 0.30;
- Coefficient on mother’s education for daughters: 0.31;
- Coefficient on father’s education for daughters: 0.26.
Heineck and Riphahn (2007) use four German cohorts of size 9 years from the SEOP to estimate a multinomial logit model with three categories of education: basic, middle, and advanced. They find:

- Child’s probability of achieving only basic education has effects 0.06, -0.27, and -0.35 for parents having education levels missing, middle, and advanced, respectively.
- Child’s probability of achieving only middle education has effects -0.05, 0.00, and -0.145 for parents having education levels missing, middle, and advanced, respectively.
- Child’s probability of achieving advanced education has effects -0.014, 0.27, and 0.50 for parents having education levels missing, middle, and advanced, respectively.
- Cohort effects are negative relative to earliest cohort for probability of child having only basic education, but positive for probability of child having middle or advanced education.
Hertz, et al. (2007) use many cohorts for many countries and regions to estimate the OLS coefficients and normalized coefficients for 44 countries and regions. They use averages across cohorts as well as regressions with cohort dummies. They find:

- USA has regression coefficient of 0.46 and correlation of 0.46;
- lowest intergenerational mobility in education is for Latin American countries, with a regional average of 0.60, compared to regional averages between 0.34 and 0.41 for the other regions;
- highest mobility is found in the Nordic region, with regional average of 0.34.
I construct a Great Gatsby Curve of the correlations reported by Hertz, et al. (2007) regressed on after tax-and-transfer Gini coefficients:

\[ \text{Intergen. Correlation of Schooling} = \beta \times \text{Gini Coefficient, After Taxes and Transfers (CIA 2014)} + \epsilon \]

where \( \beta = 0.48 \) and the \( p \)-value is 0.007.

![Graph showing the relationship between Gini coefficient and intergenerational correlation of schooling. The dashed line represents the OLS slope of 0.48 with a p-value of 0.007.](image-url)
More Recent and International Literature

Checchi, et al (2008) find using Italy’s SHIW data, which asks questions similar to OCG, regression and normalized regression coefficients for 13 generations of size 4 years each:

- father’s years of schooling on child’s (regression coefficient): 0.65 in 1910-14 to 0.36 since 1970, with generally decreasing estimates over time;
- father’s years of schooling on child’s (normalized regression coefficient): 0.57 in 1910-14 to 0.44 since 1970, with generally decreasing estimates over time;
- mother’s years of schooling on child’s (regression coefficient): 0.81 in 1910-14 to 0.35 since 1970, with generally decreasing estimates over time;
- mother’s years of schooling on child’s (normalized regression coefficient): 0.53 in 1910-14 to 0.41 since 1970, with generally decreasing estimates over time.

Note: the effects for mothers and fathers are estimated separately.
Chavelier, Denny, and McMahon (2009) regress parent’s IALS score (1-5 scale) on child’s for Europe and the USA. They find effects ranging from 0.22 in Switzerland to 0.54 in Germany, with the USA at 0.36. It is unclear what these coefficients measure.
Lindahl, Palme, Massih, and Sjogren (2012) use four generations of Swedish data to estimate the effects of parents’, grandparents’, and great-grandparents’ years of schooling on child’s years of schooling. They find regression coefficients:

- on the fourth generation of children of 0.15, 0.13, and 0.30 for great-grandparents, grandparents, and parents, respectively;
- on the third generation of children of 0.38 and 0.28 for grandparents and parents, respectively; and,
- on the second generation of children of 0.61 for parents.
More Recent and International Literature

Miscellaneous other estimates:

- Belzil and Hansen (2003) find in the USA the regression coefficients for father’s and mother’s years of schooling on child’s are 0.21 and 0.17, respectively, using NLSY 1979-1990 white sons.

- Farre, Klein, and Vella (2012) find in the USA the same result for sons, but effects of -0.31 and 0.01, respectively, for daughters.

- Bauer and Riphahn (2006) estimate the effects on the probability of child attaining high secondary education conditional on parents’ attainment in Swiss data, finding effects of 0.12, 0.28, and 0.58 for parents’ attainment low, middle, and high, respectively.
Conclusions

How mobile is the USA?

- Early estimates based on OCG placed the regression coefficient around 0.20, and the correlation around 0.45.
- However, estimates vary depending on mother’s or father’s education and on daughter’s or son’s education. Fathers tend to have greater effects on sons, and mothers tend to have greater effects on daughters.
- Recent USA estimates vary widely. For example, estimated effects of father’s years of schooling on daughter’s are as low as -0.31 in NLSY and as high as 0.26 in PSID, while the effect of father’s years on son’s ranges from 0.27 in PSID to 0.46 in ISSP category. For what it’s worth, these are all statistically significant.
Conclusions

Have countries become less mobile in education over time?

- Very unclear if the USA has become less mobile, depends on which data set is used and which parent/child combination is used.
- Italy appears to have become more mobile, but still one of the least mobile countries, while Germany has become less mobile.
- Sweden shows high mobility not only from parents to children but also from grandparents and great-grandparents. This raises questions about what is being missed in the various other studies that do not control for older generations.
References

References


