Intergenerational mobility: Some theory and empirical results.

Valentino Dardanoni

Juanary 2009

Social Mobility is not a very new topic....

You are, all of you in this community, brothers. But when god fashioned you, he added gold in the composition of those of you who are qualified to be Rulers (which is why their prestige is greatest); he put silver in the Auxiliaries, and iron and bronze in the farmers and other workers. Now since you are all of the same stock, though your children will commonly resemble their parents, occasionally a silver child will be born of golden parents, or a golden child of silver parents, and so on.

If one of the Rulers' own children has traces of bronze or iron in its make-up, they must harden their hearts, assign it its proper value, and degrade it to the ranks of the industrial and agricultural class where it properly belongs: similarly, if a child of this class is born with gold or silver in its nature, they will promote it appropriately to be a Guardian or an Auxiliary. And this they must do because there is a prophecy that the State will be ruined when it has Guardians of silver or bronze.

Plato, *The Republic*, III, 415

- At least since Plato social mobility has been deemed desirable since it brings *efficiency* and *equity*, and is closely linked to the concept of *equality of opportunity.*
- Measuring mobility is complex; some empirical results on mobility comparisons show this very clearly.
- In this presentation I will concentrate on some empirical models of intergenerational mobility with two applications: a test of equality of opportunity, and an estimation of intergenerational schooling transmission.

Galton and Becker

- Sir Francis Galton in 1886 argued that expected *stature* of the child is a weighted average between mean stature in the population and stature of parents.

$$E(S_t) = (1 - \beta)\mu + \beta S_{t-1}$$

- Stature then follows an AR(1) process

$$S_t = \alpha + \beta S_{t-1} + \epsilon_t$$

- So, stature *regresses to the mean* and β is the correlation coefficient which measures (im)mobility.
- Galton invented regression analysis, and more than a century later his analysis is the workhorse of 99% empirical models of social mobility in economics. In fact, Galton's analysis was more sophisticated...

- A theoretical justification of Galton's model has been provided by Gary Becker, incorporating elements of parents' choice.
- In Becker-Tomes's (JPE, 79 and JLabEcon, 86), income of the father Y^p is divided into own consumption C^p and investment for human capital for the child I^p:

$$Y^p = C^p + I^p$$

- The income of the son Y^c is then

$$Y^c = (1 + r)I^p + E^c$$

where r is the return on the investment and E^c is 'everything else' that influence earnings.

- Father chooses investment I^p to maximize a Cobb-Douglas $U(C^p, Y^c)$, income for the son is

$$Y^c = \beta Y^p + \alpha E^c$$

  if Y^p and E^c are orthogonal, β is the intergenerational correlation in earnings.
- Typically E^c and Y^p are not orthogonal; divide E into

$$E^c = e^c + u^c$$

  where e^c is "nature and nurture" inherited from father and u "pure market luck".

- Further, if we assume that inherited skills follow an AR(1) process

$$e_t = \lambda e_{t-1} + v$$

we get a reduced form

$$Y^c = \alpha + \beta Y^p + \gamma e^p + \epsilon$$

- Becker's model is thus observationally equivalent to Galton's model (with additional control for e^p). Mulligan (JPE, 99) shows how auxiliary asumptions can help discriminate between the two models.

How large is intergen. correlation?

- Mulligan surveys intergenerational (father-son) correlation estimates from various conutries at various times shows substantial persistence:
- Years of schooling (8 studies):
  range .14-.45, avg. .29
- Log earnings or wages (16 studies):
  range .11-.59, avg. .34
- Log family wealth (9 studies):
  range .27-.76, avg. .50
- Log family consumption (2 studies)
  range .59-.77 avg. .68

Stochastic monotonicity

Galton-Becker-Tomes describe the *conditional mean* of Y^c; does not consider the whole bivariate distribution.

- The question: “Is it better to have a rich father?” can be formalized in terms of *stochastic monotonicity*: $\forall y^c,\ F_{Y^c|Y^p}(y^c \mid y^p) \leq F_{Y^c|Y^p}(y^c \mid y^{p'})$ whenever $y^p \geq y^{p'}$
- Lee, Linton and Whang (Ecma, 2009) propose a nonparameric test for the hypothesis of stochastic monotonicity with continuous Y^p, Y^c.
- For discrete Y^p, Y^c, the nonparametric test of Dardanoni and Forcina (JASA, 1998) can be used.
- One advantage of the discrete approach is that it can be easily extended incorporating control covariates.

Unconditional stochastic monotonicity

- Dardanoni, Fiorini and Forcina (2008) apply the DF nonparametric test of stochastic monotonicity to a sample of 149 6×6 social mobility matrices, gathered for 35 countries by Ganzeboom, Luijkx, and Treiman (Research in Soc. Strat. and Mob., 1989).
- The distribution of the LR test statistic is *chi-bar-squared* and conservative α critical values can be found by solving the equation

$$\sum_{i=0}^{(k-1)^2} \binom{(k-1)^2}{i} 2^{-(k-1)^2} Pr[\chi_i^2 = c] = \alpha.$$

where k is the number of social classes.

It is better to have a rich father....

- Out of 149 tables, stochastic monotonicity is rejected at the 99% significance level only for Hungary 1963, Philippines 1968, Poland 1972 and Spain 1975. In addition, the monotonicity hypothesis is rejected at the 95% level for Hungary 1973 and 1983 and India 1963c.
- Thus, it appears that monotonicity of the intergenerational transmission mechanism can generally be considered as an assumption supported by the real world.
- This may have interesting implications for mobility measurement.

Cond. stoch. monotonicity and *EoP*

- Unconditional stochastic monotonicity does not necessarily have an equality of opportunity interpretation.
- *EoP* holds in a society if individuals' chances to succeed depend only on their own efforts; what is contentious, however, is what constitutes effort and circumstances.
- Dardanoni, Fields, Roemer and Sanchez-Puerta (in book, 2006) describe four channels through which parents affect child's status: social connections, formation of social beliefs and skills, transmission of native ability and instillation of preferences and aspirations.

- Finding unconditional stochastic monotonicity contradicts the most demanding version of *EoP*, where all those channels are circumstances out of an individual's control.
- Less stringent notions of *EoP* allow for some of those channels to be influenced by the offsprings. In turn this requires independence conditional on appropriately selected covariates.
- We develop a formal inferential procedure for testing for conditional stochastic monotonicity, and we apply it to the UK NCDS survey.

Testing cond. stochastic monotonicity

- Let X and Y denote father's and son's class, and let $\boldsymbol{z}$ be a vector of covariates which may affect the joint distribution $\boldsymbol{\pi}(\boldsymbol{z})$.
- Our approach models the effect of covariates by a suitable link function, which maps $\boldsymbol{\pi}(\boldsymbol{z})$ into a set of parameters which have monotonicity interpretation, and a regression model.
- Our parameters are *Local-Global Odds Ratios*

$$\tau_{ij}(\boldsymbol{z}) = \log\left[\frac{P(X=i, Y\leq j \mid \boldsymbol{z})P(X=i+1, Y>j \mid \boldsymbol{z})}{P(X=i, Y>j \mid \boldsymbol{z})P(X=i+1, Y\leq j \mid \boldsymbol{z})}\right].$$

- It can be shown that $\boldsymbol{\pi}(\boldsymbol{z})$ is monotone if and only if the set of $(k-1)^2$ LG-log odds ratios are nonnegative.

- Collect now the fathers and sons marginal parameters and the association parameters into the vector

$$\boldsymbol{\lambda}(\boldsymbol{z}) = [\boldsymbol{\rho}(\boldsymbol{z})^{'}, \boldsymbol{\xi}(\boldsymbol{z})^{'}, \boldsymbol{\tau}(\boldsymbol{z})^{'}]^{'}$$

- It can be shown that the mapping from $\boldsymbol{\pi}(\boldsymbol{z})$ to $\boldsymbol{\lambda}(\boldsymbol{z})$ is invertible and differentiable; *no parametric assumptions are made on the conditional distribution* $\boldsymbol{\pi}(\boldsymbol{z})$.
- The parametric structure comes from setting a linear regression model

$$\begin{aligned}
\rho_i &= \alpha_i^X + \boldsymbol{z}^{'}_X \boldsymbol{\beta}_i^X, \; i = 1, \ldots, k-1 \\
\xi_j &= \alpha_j^Y + \boldsymbol{z}^{'}_Y \boldsymbol{\beta}_j^Y, \; j = 1, \ldots, k-1 \\
\tau_{ij} &= \alpha_{ij}^{XY} + \boldsymbol{z}^{'}_{XY} \boldsymbol{\beta}_{ij}^{XY}, \; i, j = 1, \ldots, k-1
\end{aligned}$$

Hypotheses

- The hypothesis of conditional stochastic monotonicity can be expressed as appropriate linear inequality constraints:

$$\mathcal{H}_1 : \tau_{ij} = \alpha_{ij}^{XY} + \boldsymbol{z}'_{XY}\boldsymbol{\beta}_{ij}^{XY} \geq 0 \;\forall\; \boldsymbol{z}_{XY};\; i,j = 1,\ldots,k-1.$$

  Notice that in typical applications many inequalities are likely to be redundant. For example, if covariates are continuous there are as many constraints as sample points.

- Equality of opportunities can be written as

$$\mathcal{H}_0 : \tau_{ij} = \alpha_{ij}^{XY} + \boldsymbol{z}'_{XY}\boldsymbol{\beta}_{ij}^{XY} = 0 \;\forall\; \boldsymbol{z}_{XY};\; i,j = 1,\ldots,k-1;$$

Results in the British NCDS

- We use the British National Child Development Study (NCDS), an ongoing survey that originally targeted over 17,000 babies born in Britain in the week 3-9 March 1958.
- Surviving members have been surveyed on seven further occasions in order to monitor their changing health, education, social and economic circumstances.
- At the age of 7, 11 and 16 mathematics, reading and general skills tests were taken and at the age of 7 and 11 information on non - cognitive skills was also collected.

- To apply our tests we first have to find suitable variables representing socio-economic status X and Y and covariates z.
- We use 3 social and 3 wage classes, and as control covariates, we use cognitive and non-cognitive skills, educational attainment (at age 23) and father's age.
- For both social and wage class, we first estimate the urestricted, the equality of opportunities and the stochastic monotonicity models by ML.
- The LR test statistics are again shown to be asymptotically distributed as *chi-bar-square.*

Results

- Social class mobility: p-values

Row	*EoP vs SM*	*SM vs Unr*
R_{12}	0.0003	0.8450
R_{23}	0.0000	0.8284
AR	0.0000	0.9460

- Wage mobility: p-values

Row	*EoP vs SM*	*SM vs Unr*
R_{12}	0.3342	0.5574
R_{23}	0.0005	0.0701
AR	0.0007	0.3064

- Stochastic monotonicity cannot be rejected even conditional on educational achievement, cognitive and non-cognitive skills.
- There seems to be more conditional mobility in wage than in social class. However this may be due to measurement error (attenuation bias) or to sample selection.

Intergener. Education Transmission

- What is the effect of fathers' and mothers' schooling on children's education?
- Old conventional wisdom: both parents count; mothers count "a bit more".
- Challenge: this reflects correlation and not causation.
- The challenge starts with Behrman and Rosenzweig, who find that mothers' schooling has *no* effect on children's. This has generated some controversy and has important policy implications.

G-B-T revisited by Solon (1999)

- The standard model for analyzing intergenerational schooling transmission is an adaptation of the G-B-T model by Solon (in book, 1999).
- Again, parent chooses child's investment I^p to maximize $U(Y^c, C^p)$, and child's schooling is a function of parent's investment and "everything else":

$$S^c = S^c(I^p, N^c)$$

with

$$N^c = N^c(R^p, U^p, S^p)$$

with R^p being parent's child rearing ability, U^p inheritable parent's ability and S^p parent's schooling.

Reduced form

- Using a standard Mincerian specification for the income equation and appropriate functional forms for the other equations and the utility function, we get to the reduced form equation

  $$S^c = \alpha + \beta S^p + \gamma R^p + \delta U^p + \epsilon$$

  where β agains captures the intergenerational schooling transmission coefficient.
- The problem in estimating β is that without proper control for R^p and U^p the standard regression coefficient is biased.
- To identify β, three strategies are typically used in this literature.
- As an aside, notice alternative literature which considers *child*'s, not parent's rational choice

Twins: BR, AER 2002-2005

- How do we estimate β, since the schooling equation contains unobservables?
- Berhman and Rosenzweig use MZ twins.
- Let $(S_t^c, S_t^p, U_t^p, R_t^p)$ refer to MZ twins $t = 1, 2$.
- IF $U_1^p = U_2^p$ and $R_1^p = R_2^p$, take differences, and thus

$$(S_1^c - S_2^c) = \beta(S_1^p - S_2^p) + \phi$$

  and estimate β.
- They find no effect of mothers' schooling on children's education, and a substantial effect of fathers' schooling. Explanation?
- Problem: do $U_1^p = U_2^p$ and $R_1^p = R_2^p$?

Adoptees: Plug, AER 2004

- Plug uses adoptees. The advantage is that in the equation $S^c = \alpha + \beta S^p + \gamma R^p + \delta U^p + \epsilon$, δ is equal to zero by assumption if it measures genetic transmission.
- Problem: Do better educated parents adopt more endowed child? Is adoption process random?
- Plug finds little effect of mothers' schooling on children's education.

IV: Black, Devereux, Salvanes AER 2006

- Black, Devereux, Salvanes use IV estimation using Norwegian data.
- They use a compulsory schooling law change, implemented in different municipalities between 1960 and 1972.
- The reform provides variation in parental education that is exogenous to parental endowments, and municipalities can be used as instrument.
- They find no effect of fathers' schooling on children's education, while mothers' schooling slightly affects son's education.
- Problem: use of IV.... municipalities may not be exogeneous...

Estimating β...

- It takes two parents to make one child.... Controlling for U^m or U^f separately implies problems with assortative mating.
- Controlling for parents' endowments requires much care and ingenuity; estimates may be quite sensitive to the key assumptions made.
- Other studies have used the three strategies to identify β with different datasets. Results are generally conlficting between methods.
- Holmlund, Lindahl and Plug (2008) argue that the different results obtained by the three methods are not due to the different data sets used, but by the different identification strategies, by appliying the the three different methods to the same Swedish data set.

Direct and indirect causal effects...

- Sociologists (Boudon, 1973, Erickson, Goldthorpe, Jackson, Yaish and Cox, PNAS, 2005) argue that parents' background has two effects of child's outcome:
  1. A *primary effect*: the effect of family background on academic ability
  2. A *secondary effect*: the effect of family background on educational attainment *given* academic ability.
- Estimated secondary effects are found important, accounting on average about a quarter of the total effect.
- In the language of causality analysis, primary and secondary effects are called *indirect* and *direct*.

Causal effects in the linear model

- The benchmark equation

$$S^c = \alpha + \beta S^p + \gamma R^p + \delta U^p + \epsilon$$

can be decomposed as

$$U^c = d + eS^p + fR^p + gU^p + \eta$$

and

$$S^c = a + bS^p + cU^c + \psi$$

- In this framework, total "causal effect"of S^p on S^c is reduced form coefficient $\beta = b + ce$, indirect effect is ce and direct effect is b.
- Note this decomposition does not hold in nonlinear model.

The technology of skills formation...

- Recently, Heckman and collaborators (e.g. Cunha and Heckman, AER 2007, Cunha, Heckman and Shennack, Ecma 2009) are actively investigating the technology of skills formation, how skills evolve through time, and how parental background affects unobservable abilities (cf equation for U^c above).
- In particular, they are concerned with the difference between early and late intervention, and the measure of the dynamic complementarity of skills.
- They find very early intervention is much more cost effective.

Direct causal effect: DFM (2008)

- We estimate the direct causal effect (recall coefficient b in equation above).
- Thus, we have to control for U^c and not U^p.
- Interest: for understanding channels of equality of opportunity...
- As child's schooling outcome, we use a binary variable of attainment of a given degree (in our application, English O-Level certification).
- Difference between *attainment* and *potential ability*; why should parents' schooling enter the equation for S^c after knowing child's unobservable skills?
- Possible answer: *role effects*, emulation, job market prospects, info on the value of education, etc.

How we do it...

We identify U^c by:

1. Using the very rich UK NCDS dataset, which contains data on educational attainment and early cognitive and noncognitive skills.
2. Using marginal modelling techniques applied to finite mixture models.

The data

- Students at 16 take O-levels exams; if a sufficient number of O-Levels is passed the child is allowed to access the next level of education (A-Levels).
- Knowing whether the subject has passed enough O-Level exams (we construct a binary variable OL), and knowing cognitive and non cognitive test scores, we can differentiate between scholastic *attainment* and *potential ability*.
- Our sample contains 2627 sons and 2568 daugthers; about fifty percent of subjects achieve OL.

- We construct binary variables EM, LM, ER, LR, ENP, ENS which measure early (7 and 11) and late (16) math and reading ability, and early personal and social noncognitive skills.
- As for family background variables we have:

1. Parents' schooling: age when left education, in years, denoted $\boldsymbol{s}$.
2. Parents' interest in their child's education (proxies for R^f and R^m), denoted $\boldsymbol{r}$.

The model we estimate

- We estimate the following multivariate nonlinear regression system:

$$
\begin{aligned}
Pr(OL = 1 \mid u, \boldsymbol{s}) &= \Lambda\Big(a_{OL}(u) + \boldsymbol{s}^{'}\boldsymbol{\beta}_{OL}\Big) \\
Pr(EM = 1 \mid u) &= \Lambda\Big(a_{EM}(u)\Big) \\
Pr(LM = 1 \mid em, u) &= \Lambda\Big(a_{LM}(u) + b_{LM}em\Big) \\
Pr(ER = 1 \mid u) &= \Lambda\Big(a_{ER}(u)\Big) \\
Pr(LR = 1 \mid er, u) &= \Lambda\Big(a_{LR}(u) + b_{LR}er\Big) \\
Pr(ENP = 1 \mid u) &= \Lambda\Big(a_{ENP}(u)\Big) \\
Pr(ENS = 1 \mid u) &= \Lambda\Big(a_{ENS}(u)\Big) \\
Pr(U = u \mid \boldsymbol{s}, \boldsymbol{r}) &= \Gamma\Big(a_U(u) + \boldsymbol{x}^{'}\boldsymbol{\beta}_U(u)\Big)
\end{aligned}
$$

- The model can be see as a semiparametric random effect model, with random effects identified by means of auxiliary regressions.
- Parameters are estimated by EM algorithm.

Results

Simulations reveal parameters are well identifed. Actual estimates have small se. Model selection criteria suggests 3 latent types. Estimated parameters reveal:

- Monotonicity and unidimensionality of U: LR test equal 0.0036 for males and 0.0042 for females.
- High differences in probabilities of success in all test scores for the 3 types.
- Strongly significant recursive effects, even after controlling for U.
- Strong positive association between parents' background and U (Caveat: this cannot be interpreted as a causal relation...).

Intercepts in OL-equation

	Daughters		Sons	
	coeff	se	coeff	se
$a_{OL}(U=0)$	-1.8736	0.1476	-2.5742	0.1918
$a_{OL}(U=1)$	0.5589	0.1269	0.2164	0.1166
$a_{OL}(U=2)$	2.4019	0.2234	2.2798	0.2372

Thus, very strong ability random effects on OL attainment; for example, a logit of 2 implies $p \sim .88$ and a logit of -2 implies $p \sim .12$, while a logit equal 0 implies $p = .5$.

Parents' schooling in OL-equation

	Daughters		Sons	
	coeff	se	coeff	se
fs	-0.0173	0.0454	0.1150	0.0488
ms	0.0831	0.0532	-0.0363	0.0549

Thus significant father/son and weak mother/daugther effects..

Direct causal effect

- To get a quantitative feeling of these coefficients, we consider the effect on OL attainment of increasing each parent education by three full years of schooling:

$$\begin{aligned}\delta(u)_{S^i} &= Pr(OL = 1 \mid S^j = \mu^j, S^i = \mu^i + 3, U = u) - \\ &\quad Pr(OL = 1 \mid S^j = \mu^j, S^i = \mu^i, U = u)\end{aligned}$$

	Daughters		Sons	
	$\delta(u)_{Sf}$	se	$\delta(u)_{Sf}$	se
U=0	-0.0059	0.0098	0.0263	0.0129
U=1	-0.0121	0.0380	0.0829	0.0392
U=2	-0.0040	0.0137	0.0253	0.0122
	$\delta(u)_{Sm}$	se	$\delta(u)_{Sm}$	se
U=0	0.0315	0.0248	-0.0068	0.0129
U=1	0.0555	0.0403	-0.0270	0.0442
U=2	0.0171	0.0116	-0.0096	0.0156

Pooled sons and daughters sample

- We have also estimated the previous model on the pooled sample of sons and daughters.
- Pooled sample estimates are approximately equal to the average of the corresponding estimates for daughters and sons:

	coeff	se
fs	0.0452	0.0380
ms	0.0253	0.0403

- Thus, no significant direct effect of fathers' and mothers' on children's.

Females and Attainment..

- To help understand our result, we check whether there is a gender specific bias in grading and achievements, conditional on true potential ability.
- Bias is documented for many countries with 2003 PISA data set by Dardanoni, Modica and Pennisi 2008.
- In a simple logit of OL on test scores we estimate a very significant female dummy coefficient equal to 0.6129 with se equal to 0.0786. This translates, for an individual with average test scores, in a probability of achieving OL equal to 0.612 for females and 0.461 for males.

Conclusions

- In the pooled sample we find no direct effect of parents' education.
- Considering sons and daughters subsamples separately, after controlling for unobs. het., the predominance of the paternal figure is entirely confined to boys. Girls appear to be only influenced by mothers, but the effect is not significant.
- We DO NOT:

  1. Understand how children get to be in different unobserved ability types, and what can be done about it (Heckman et al...);
  2. Estimate the *total effect* of parents education on children's

That's it

THANK YOU