sociological science

Supplement to:

Breen, Richard, and John Ermisch. 2024. "The Effects of Social Mobility." Sociological Science 11: 467-488.

Supplemental appendixes to Breen and Ermisch, *The Effects of Social Mobility*

Appendix A: The dependence of "mobility effects" estimates on the coding scheme

Luo (2022) presents an approach to the definition and estimation of mobility effects based on the SAM (as shown in our equation 1). Luo defines mobility effects, which we label c_{jk}, as $c_{jk} = \gamma_{jk} - \gamma_{jj}$: that is, c_{jk} is the difference between the interaction parameters for a cell representing mobility from a given origin (where $k \neq j$) and that for the cell representing immobility for the same origin. The problem of the lack of identification of the full set of interactions is addressed by the use of effect coding which constrains the interactions to sum to zero across rows and columns as shown below for the 3 by 3 case. But this means that Luo's mobility effects, c_{jk} depend on the coding scheme chosen.

Table A1: Constraints with effect coding

γ_{11}	γ_{12}	$\gamma_{13} = -(\gamma_{11} + \gamma_{12})$
γ_{21}	γ_{22}	$\gamma_{23} = -(\gamma_{21} + \gamma_{22})$
$\gamma_{31} = -(\gamma_{11} + \gamma_{21})$	$\gamma_{32} = -(\gamma_{12} + \gamma_{11})$	$\gamma_{33} = -(\gamma_{13} + \gamma_{23}) = -(\gamma_{31} + \gamma_{32})$

Effect coding is simply a way of constraining the interaction terms and it inevitably leads to constraints on the mobility effects. For example, for the 3 by 3 case we have

$$c_{13} + c_{23} = c_{21} + c_{31} = c_{12} + c_{32}$$

It is not clear why the mobility effects from origins 1 and 2 should be constrained by mobility effects from origin 3.

Calculation of γ_{ij} and c_{jk} *from table of means*

Effect coding

The effect coding scheme leads to simple expressions for the estimates of γ_{ij} ($\hat{\gamma}_{ij}$) in terms of the means in the data. Define μ_{ij} as the deviation of the mean of cell *i,j* from the overall mean (μ), and, for our 3x3 example, define $\mu_{i\cdot} = \sum_{j=1}^{3} \frac{\mu_{ij}}{3}$ as the column mean (net of the grand mean) and $\mu_{\cdot j} = \sum_{i=1}^{3} \frac{\mu_{ij}}{3}$ as the row mean (net of the grand mean), noting that $\sum_{i=1}^{3} \mu_{i\cdot} = \sum_{j=1}^{3} \mu_{\cdot j} = 0$.

Then $\hat{\alpha}_1 = \mu_{1\cdot}$, $\hat{\alpha}_2 = \mu_{2\cdot}$, $\hat{\beta}_1 = \mu_{\cdot 1}$, $\hat{\beta}_2 = \mu_{\cdot 2}$, and

$\hat{\gamma}_{11} = \mu_{11} - \hat{\alpha}_1 - \hat{\beta}_1$ $(= \mu_{11} - \mu_{1\cdot} - \mu_{\cdot 1})$

$\hat{\gamma}_{12} = \mu_{12} - \hat{\alpha}_1 - \hat{\beta}_2$ $(= \mu_{12} - \mu_{1\cdot} - \mu_{\cdot 2})$

$\hat{\gamma}_{21} = \mu_{21} - \hat{\alpha}_2 - \hat{\beta}_1$ $(= \mu_{21} - \mu_{2\cdot} - \mu_{\cdot 1})$

$\hat{\gamma}_{22} = \mu_{22} - \hat{\alpha}_2 - \hat{\beta}_2$ $(= \mu_{22} - \mu_{2\cdot} - \mu_{\cdot 2})$

The remaining $\hat{\gamma}_{ij}$ can be derived from these estimates by the adding-up constraints (e.g. $\hat{\gamma}_{31} = -(\hat{\gamma}_{11} + \hat{\gamma}_{21})$.

The estimated mobility contrasts $\hat{c}_{jk} = \hat{\gamma}_{ij} - \hat{\gamma}_{ii}$ are derived from the $\hat{\gamma}_{ij}$. For example,

$\hat{c}_{12} = \mu_{12} - \mu_{11} + \hat{\beta}_1 - \hat{\beta}_2$ $(= \mu_{12} - \mu_{11} + \mu_{\cdot 1} - \mu_{\cdot 2})$

$\hat{c}_{21} = \mu_{21} - \mu_{22} + \hat{\beta}_2 - \hat{\beta}_1$ $(= \mu_{21} - \mu_{22} + \mu_{\cdot 2} - \mu_{\cdot 1})$

$\hat{c}_{13} = \mu_{13} - \mu_{11} + \hat{\beta}_1 - \hat{\beta}_3$ $(= \mu_{\cdot 2} + 2\mu_{\cdot 1} + 3\mu_{1\cdot} - 2\mu_{11} - \mu_{12})$

Dummy variable coding

Suppose that we constrain $\alpha_2 = \beta_2 = \gamma_{12} = \gamma_{21} = \gamma_{23} = \gamma_{32} = \gamma_{22} = 0$. Define m_{ij} as the deviation of the mean of cell *I,j* from the mean of cell *2,2*.

Then $\hat{\alpha}_1 = m_{12}$, $\hat{\alpha}_3 = m_{32}$, $\hat{\beta}_1 = m_{21}$, $\hat{\beta}_3 = m_{23}$, and

$\hat{\gamma}_{11} = m_{11} - \hat{\alpha}_1 - \hat{\beta}_1\ (= m_{11} - m_{12} - m_{21})$

$\hat{\gamma}_{13} = m_{13} - \hat{\alpha}_1 - \hat{\beta}_3\ (= m_{13} - m_{12} - m_{23})$

$\hat{\gamma}_{31} = m_{31} - \hat{\alpha}_3 - \hat{\beta}_1\ (= m_{31} - m_{32} - m_{21})$

$\hat{\gamma}_{33} = m_{33} - \hat{\alpha}_3 - \hat{\beta}_3\ (= m_{33} - m_{32} - m_{23})$.

The estimated mobility contrasts $\hat{c}_{jk} = \hat{\gamma}_{ij} - \hat{\gamma}_{ii}$ are:

$\hat{c}_{13} = m_{13} - m_{11} + \hat{\beta}_1 - \hat{\beta}_3\ (= m_{13} - m_{11} + m_{21} - m_{23})$

$\hat{c}_{31} = m_{31} - m_{33} + \hat{\beta}_3 - \hat{\beta}_1\ (= m_{31} - m_{33} + m_{23} - m_{21}$

$\hat{c}_{32} = -\hat{\gamma}_{33}$

$\hat{c}_{12} = -\hat{\gamma}_{11}$ and $\hat{c}_{21} = 0 = \hat{c}_{32}$.

Under the approach we present in this paper, the estimated causal effect of moving from origin *i* to destination *j* is $\hat{\beta}_j - \hat{\beta}_i + \hat{c}_{ij}$. When origins and destinations are randomly assigned this can be estimated by ordinary least squares. Substituting from above, the causal effect is $m_{ij} - m_{ii}$ with dummy variable coding and $\mu_{ij} - \mu_{ii}$ with effect coding but $m_{ij} - m_{ii} = \mu_{ij} - \mu_{ii}$ because both are just differences in the cell means of *I,j* and *i,i*. In other words, the different coding schemes produce different values of $\hat{\beta}_j - \hat{\beta}_i$ and therefore of $\hat{c}_{ij}$, but not different estimates of causal effects of moving from origin *i* to destination *j*.

Example

To illustrate the impact of the different coding schemes on estimates of $\hat{c}_{ij}$, data on family size and social class from Berent (1952) was used: these data were also employed by Duncan (1966) in his seminal article. We have combined the top two social classes in Berent's scheme into class 1 with classes 2 and 3 being the same as Berent's. The data are shown in Table A2. Downward mobility occurs above the diagonal and upward mobility below it. Upward mobility results in lower fertility, and downward higher fertility. Mobility contrast estimates under the two coding schemes are shown in Table A3 and the underlying parameter estimates in Table A4.

Table A2: Average Family size and social class

Father/Son	1	2	3
1 (top)	2.01	2.44	2.83
2	2.13	2.67	3.69
3	1.98	3.22	3.68

Table A3: Mobility contrast estimates under the two coding schemes

	Effect	Dummy
$\hat{c}_{12}$	-0.31	-0.12
$\hat{c}_{13}$	-0.54	-0.74
$\hat{c}_{21}$	0.19	0
$\hat{c}_{23}$	0.39	0
$\hat{c}_{31}$	-0.34	-0.13
$\hat{c}_{32}$	0.17	0.56

Although the mobility contrasts are in the same direction in the two coding schemes they differ substantially in magnitude in some cases, and hypothesis tests on individual contrasts may produce different results in smaller samples. In terms of substantive findings, if we were to interpret the contrasts as 'mobility effects', fertility is lower for the downwardly mobile from class 1 and the upwardly mobile from class 3 to 1, but higher for the upwardly mobile from class 3 to 2. These 'mobility effects' clash with the causal mobility effects in two of these three instances: fertility is higher for the downwardly mobile from class 1 to 3 and lower for the upwardly mobile from class 3 to 2.

Table A4: Parameter estimates under the two coding schemes

	Effect	Dummy
$\hat{\mu}$	2.74	2.67
$\hat{\beta}_1$	-0.70	-0.54
$\hat{\beta}_2$	0.04	0
$\hat{\beta}_3$	0.66	1.02
$\hat{\alpha}_1$	-0.31	-0.23
$\hat{\alpha}_2$	0.09	0
$\hat{\alpha}_3$	0.22	0.55
$\hat{\gamma}_{11}$	0.28	0.12
$\hat{\gamma}_{12}$	-0.03	0
$\hat{\gamma}_{13}$	-0.26	-0.62
$\hat{\gamma}_{21}$	0.00	0
$\hat{\gamma}_{22}$	-0.20	0
$\hat{\gamma}_{23}$	0.20	0
$\hat{\gamma}_{31}$	-0.28	-0.69
$\hat{\gamma}_{32}$	0.22	0
$\hat{\gamma}_{33}$	0.06	-0.56

Appendix B: Estimates under Conditional Independence Assumption

In order for propensity score weighting or matching to produce a consistent estimate of the average causal effect of social mobility from origin *j* to destination *k* on those moving to destination k, $ATT_{jkj} = E\ (Y_i(k) - Y_i(j)|O = j, D = k)$, we need a Conditional Independence Assumption (CIA) analogous to the Propensity Score Theorem (Angrist and Pischke 2009: 80). For those from origin *j*, let $p_{jk}(X_i)$ be the probability that $D_i = k$ which is the probability of mobility from *j* to *k*. Then the CIA theorem states that if $Y_i(D)$ is orthogonal to D_i given X, then $Y_i(D)$ is also orthogonal to D_i given $p_{jk}(X_i)$. The theorem holds for the unknown true propensity score: [1] what follows depends on how well we can estimate it.

Define $I_i = 1 \text{ for } D_i = k \text{ and } I_i = 0 \text{ for } D_i = j.$ Then $p_{jk}(X_i) = E(I_i|O = j, D_i = k, X_i)$ and the inverse propensity score weighted estimate of ATT_{jkj} is

$$(1/N)\sum_{i=1}^{N} \frac{Y_i I_i}{p_{jk}(X_i)} - \frac{(1 - I_i)Y_i}{(1 - p_{jk}(X_i))}$$

where N is the sample size, which could be restricted to observations with common support. This estimator for destination k and origin j replaces the φ_{jk} parameters in equations (7) for $k \neq j$ and from these we can derive φ_{j0}, from which we can compute the SAM parameters, as discussed in the paper after equations (7) and (8).

[1] Although the CIA holds for the true propensity score in the population, the propensity score is unknown to the analyst who must estimate it using sample data. If the estimated propensity score is a consistent estimator of the true score then conditional independence will hold in expectation, just as, with randomization, the treatment variable will be orthogonal to unobserved confounders in expectation.

In section 5 of the paper, the estimate of the probability of mobility of each type and of immobility is derived from a multinomial logit model to predict destination, conditional on origin j, from which we obtain estimates of the probability that the destination is k, $\hat{P}_{jk}$, k=1,2,3. Thus, the estimate of the propensity score for movement from origin j to destination k is $\hat{p}_{jk}(X_i) = \frac{\hat{P}_{jk}}{\hat{P}_{jk}+\hat{P}_{jj}} | X_i, j \neq k$.

References:

Angrist , Joshua and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics*. Princeton, NJ: Princeton University Press.