ONLINE SUPPLEMENT

Song, Xi and Robert D. Mare. 2015. "Retrospective Versus Prospective Approaches to the Study of Intergenerational Social Mobility." *Sociological Methods and Research* 44(4): 555-584.

Appendix A: Simulation Details

This appendix provides the details for the simulation examples. For the two-generation model, we assume that a man's fertility depends on his own socioeconomic position, and his socioeconomic position depends on only his father's position. There is no lagged effect from the grandfather in both the fertility and mobility equations. The data are generated in the following order according to the specified probability models:

1.1 The exogenous variable U_1 for the fathers' generation is drawn from a standard normal distribution $U_1 \sim N(0, 1)$.

1.2. We then generate the father's position Y_1 (1 vs. 2) for each of the 10,000 subjects:

$$\text{logit}(P[Y_1 = 2|U_1]) = \alpha_0 + \alpha_1 \cdot U_1 = \log\left(\frac{0.3}{0.7}\right) + \log(2) \cdot U_1 \quad \text{(A-1)}$$

1.3. The conditional distribution of a father's fertility given Y_1 follows a Poisson distribution with the mean of the fertility satisfies the equation below.

$$E(F_1|Y_1) = \exp\left(\beta_0 + \beta_1 \cdot (Y_1 - \overline{Y_1})\right) = \exp\left(\log(1.1) + 0.6 \cdot (Y_1 - \overline{Y_1})\right) \quad \text{(A-2)}$$

We then generate a dichotomous variable D_1 based on F_1, where $D_1 = 1$ if $F_1 > 0$ and $D_1 = 0$ if $F_1 = 0$.

1.4. The conditional distribution of a son's variable U_2 given U_1 and D_1 is drawn from a normal distribution, where the standard deviation is fixed at 1 and the mean satisfies the equation below.

$$E(U_2|U_1, D_1 = 1) = \gamma_1 \cdot (U_1 - \overline{U_1}) = 0.8 \cdot (U_1 - \overline{U_1}) \quad \text{(A-3)}$$

1.5. The conditional distribution of a son's socioeconomic position Y_2 given U_2, D_1 and Y_1 follows a Bernoulli distribution.

$$\begin{aligned}\text{logit}(P[Y_2 = 2|U_2, Y_1, D_1 = 1]) &= \delta_0 + \delta_1 \cdot U_2 + \delta_2 \cdot Y_1 \\ &= \log\left(\frac{0.2}{0.8}\right) + \log(2) \cdot U_2 + \log(2.5) \cdot Y_1\end{aligned} \quad \text{(A-4)}$$

1.6. The conditional distribution of a son's fertility F_2 given Y_2, and D_1 follows a Poisson distribution with the mean of the fertility satisfies the equation below.

$$E(F_2|Y_2, D_1 = 1) = \exp\left(\beta_0 + \beta_1 \cdot (Y_2 - \overline{Y_2})\right) = \exp\left(\log(1.1) + 0.6 \cdot (Y_2 - \overline{Y_2})\right) \quad \text{(A-5)}$$

We then generate a dichotomous variable D_2 based on F_2, where $D_2 = 1$ if $F_2 > 0$ and $D_2 = 0$ if $F_2 = 0$.

In the prospective sample, all the variables F_1, F_2, D_1, D_2, Y_1, Y_2 are observed, while in the retrospective sample, we only know $F_1 > 0$ (i.e., $D_1 = 1$), F_2, D_2, Y_1, Y_2. We need to use the proportion of childless adults in the sons' generation ($D_2 = 1$) to approximate that of the fathers' generation ($D_1 = 1$) in the adjusted retrospective method.

For the three-generation model, we assume that a man's fertility depends on the socioeconomic positions and fertility of all prior generations, as well as his own socioeconomic position. In addition, we assume a man's socioeconomic position depends on the socioeconomic positions of all prior generations. We generate the data by the following steps.

2.1. The exogenous variable U_1 for the grandfathers' generation follows a standard normal distribution $U_1 \sim N(0, 1)$.

2.2. The conditional distribution of a grandfather's socioeconomic position Y_1 (1 vs. 2) given U_1, follows a Bernoulli distribution.

$$\text{logit}(P[Y_1 = 2|U_1]) = \alpha_0 + \alpha_1 \cdot U_1 = \log\left(\frac{0.3}{0.7}\right) + \log(2) \cdot U_1 \qquad \text{(A-6)}$$

2.3. The conditional distribution of a grandfather's fertility F_1 given Y_1 follows a Poisson distribution with the mean of the fertility satisfies the equation below.

$$E(F_1|Y_1) = \exp\big(\beta_0 + \beta_1 \cdot (Y_1 - \overline{Y_1})\big) = \exp\big(\log(1.1) + 0.6 \cdot (Y_1 - \overline{Y_1})\big) \qquad \text{(A-7)}$$

We generate a dichotomous variable D_1 based on F_1, where $D_1 = 1$ if $F_1 > 0$ and $D_1 = 0$ if $F_1 = 0$.

2.4. The conditional distribution of a father's variable U_2 given U_1 and D_1 follows a normal distribution, where the standard deviation is fixed at 1 and the mean satisfies the equation below.

$$E(U_2|U_1, D_1 = 1) = \gamma_1 \cdot (U_1 - \overline{U_1}) = 0.8 \cdot (U_1 - \overline{U_1}) \qquad \text{(A-8)}$$

2.5. The conditional distribution of a father's position Y_2 given U_2, F_1, D_1 and Y_1 follows a Bernoulli distribution.

$$\begin{aligned} \text{logit}(P[Y_2 = 2|U_2, Y_1, D_1 = 1]) &= \delta_0 + \delta_1 \cdot U_2 + \delta_2 \cdot Y_1 + \delta_3 \cdot F_1 \\ &= \log\left(\frac{0.2}{0.8}\right) + \log(2) \cdot U_2 + \log(2.5) \cdot Y_1 + \log(1.1) \cdot F_1 \end{aligned} \qquad \text{(A-9)}$$

2.6. The conditional distribution of a father's fertility F_2 given Y_2, Y_1, F_1 and D_1 follows a Poisson distribution with the mean of the fertility satisfies the equation below.

$$\begin{aligned} &E(F_2|Y_2, Y_1, F_1, D_1 = 1) \\ &\quad = \exp\big(\theta_0 + \theta_1 \cdot (Y_2 - \overline{Y_2}) + \theta_2 \cdot (Y_1 - \overline{Y_1}) + \theta_3 \cdot (F_1 - \overline{F_1})\big) \\ &\quad = \exp\big(\log(1.1) + 0.4 \cdot (Y_2 - \overline{Y_2}) + 0.2 \cdot (Y_1 - \overline{Y_1}) + 0.1 \cdot (F_1 - \overline{F_1})\big) \end{aligned} \qquad \text{(A-10)}$$

We generate a dichotomous variable D_2 based on F_2, where $D_2 = 1$ if $F_2 > 0$ and $D_2 = 0$ if $F_2 = 0$.

2.7. The conditional distribution of a son's variable U_3 given U_2, U_1 and D_2 follows a normal distribution, where the standard deviation is fixed at 1 and the mean satisfies the equation below.

$$E(U_3|U_2, U_1, D_2 = 1) = \pi_1 \cdot (U_2 - \overline{U_2}) + \pi_2 \cdot (U_1 - \overline{U_1}) \quad \text{(A-11)}$$
$$= 0.6 \cdot (U_2 - \overline{U_2}) + 0.2 \cdot (U_1 - \overline{U_1})$$

Note that when $D_2 = 1$, we must have $D_1 = 1$.

2.8. The conditional distribution of a son's position Y_3 given U_3, Y_2, U_2, F_2, Y_1, U_1, and D_2 follows a Bernoulli distribution.

$$\text{logit}(P[Y_3 = 2|U_3, Y_2, U_2, F_2, Y_1, U_1, D_2 = 1])$$
$$= \lambda_0 + \lambda_1 \cdot U_3 + \lambda_2 \cdot Y_2 + \lambda_3 \cdot U_2 + \lambda_4 \cdot F_2 + \lambda_5 \cdot Y_1 + \lambda_6 \cdot U_1 \quad \text{(A-12)}$$
$$= \log\left(\frac{0.15}{0.85}\right) + \log(1.8) \cdot U_3 + \log(2.0) \cdot Y_2 + \log(1.3) \cdot U_2$$
$$+ \log(1.1) \cdot F_2 + \log(1.5) \cdot Y_1 + \log\,(1.1) \cdot U_1$$

2.9. The conditional distribution of a son's fertility F_3 given Y_3, Y_2, Y_1, F_2, F_1 and D_2 follows a Poisson distribution with the mean of the fertility satisfies the equation below.

$$E(F_3|Y_3, Y_2, Y_1, F_2, F_1, D_1 = 1)$$
$$= \exp\big(\zeta_0 + \zeta_1 \cdot (Y_3 - \overline{Y_3}) + \zeta_2 \cdot (Y_2 - \overline{Y_2}) + \zeta_3 \cdot (Y_1 - \overline{Y_1}) + \zeta_4$$
$$\cdot (F_2 - \overline{F_2}) + \zeta_5 \cdot (F_1 - \overline{F_1})\big) \quad \text{(A-13)}$$
$$= \exp\big(\log(1.1) + 0.36 \cdot (Y_3 - \overline{Y_3}) + 0.20 \cdot (Y_2 - \overline{Y_2}) + 0.10$$
$$\cdot (Y_1 - \overline{Y_1}) + 0.10 \cdot (F_2 - \overline{F_2}) + 0.03 \cdot (F_1 - \overline{F_1})\big)$$