Program files:

1. STATA Programs used to obtain the data sample used by the paper from the original Add Health data, run these 3 files in order:
   variables_smoke_G10.do
   weight1.do
   sub_data_G10.do

2. MATLAB programs used to produce the estimation results:
   Run the codes under folders 1-6 to get results for Models 1-6, respectively.

3. The orders of the variables in the estimation results from the MATLAB programs are:

   3.1 Own characteristics:

   Age

   Years in School

   Male

   Black

   Asian

   Hispanic

   Other race

   Live with both parents

   Mom education less than HS

   Mom education more than HS

   Mom education missing

   Mom on Welfare

   Mom job missing

   Mom's job is Professional

   Mom other jobs

   Sport

   Age^2/10

   3.2 School dummies *(if they are included in the model)*

   3.3 Contextual effects:

Age

Years in School

Male

Black

Asian

Hispanic

Other race

Live with both parents

Mom education less than HS

Mom education more than HS

Mom education missing

Mom on Welfare

Mom job missing

Mom's job is Professional

Mom other jobs

Sport

Age^2/10

3.4 The last three are:

Endogenous effect λ (if it is included in the model)

Constant Term

Random Effect (if it is included in the model)

4. For Marginal Effects:

   4.1 Run the file *find_group.m* to locate the representative group;

   4.2 Run the *marginal_model1.m - marginal_model6.m* to obtain marginal effects for Models 1-6, respectively.

   4.3 The orders of the marginal effects for the variables from the MATLAB programs are:

      4.3.1 *Own Characteristics*

         Age

Years in School

Male

Black

Asian

Hispanic

Other race

Live with both parents

Mom education less than HS

Mom education more than HS

Mom education missing

Mom on Welfare

Mom job missing

Mom's job is Professional

Mom other jobs

Sport

4.3.2 *Contextual characteristics*

In the same order as above, except that the last one should be left out (as it is calculated for W*age^2).

4.3.3 *Endogenous effect*