# HEDG

HEALTH, ECONOMETRICS AND DATA GROUP

THE UNIVERSITY of York

WP 24/02

# Intergenerational persistence of education, smoking and birth weight: evidence from three generations

Chiara Costi; Giuseppe Migali and Eugenio Zucchelli

February 2024

http://www.york.ac.uk/economics/postgrad/herc/hedg/wps/

# Intergenerational persistence of education, smoking and birth weight: evidence from three generations

Chiara Costi[*] Giuseppe Migali[†] Eugenio Zucchelli[‡]

January 31, 2024

**Abstract**

The identification of factors affecting birth weight is a key issue in human development due to its established associations with long-term health, educational and labour-market outcomes. This paper exploits intergenerational information on three generations (grandparents-parents-children) to explore the effects of parental education and different parental smoking behaviours on birth weight. We use the intergenerational association between grandparents' education and smoking behaviour and the corresponding parental variables to aid the identification of parents' education and smoking on birth weight. We employ rich intergenerational data drawn from the US National Longitudinal Study of Adolescent to Adult Health and a flexible two-stage residual inclusion approach. We find that there is a strong intergenerational persistence of education levels and smoking behaviours across generations. Higher parental education reduces the likelihood of children's low birth weight, although the effect appears to be mainly driven by fathers. While maternal smoking during pregnancy increases the occurrence of low birth weight, especially among non-white parents, parental regular smoking (of either mothers or fathers) does not seem to greatly affect birth weight. Results are confirmed by robustness checks excluding direct effects of grandparents' smoking while in utero and using an instrumental variable for parental education.

*JEL Classification: I00, I12, I29*
*Keywords: parental education; prenatal smoking; birth weight; intergenerational persistence; Add Health.*

**Acknowledgments:** Eugenio Zucchelli gratefully acknowledges financial support from the Tómas y Valiente Fellowship funded by the Madrid Institute for Advanced Study (MIAS), Universidad Autonoma de Madrid (UAM), Spain, and grant PID2022-137819NB-I00 funded by the Spanish government. We thank attendants to seminars at RWI (University of Duisburg-Essen), ICAE (Instituto Complutense de Analisis Economico) and presentations at the 10th EuHEA (European Health Economics Association) PhD and Supervisor Conference, the 8th Portuguese National Conference on Health Economics and the 14th IWAEE (International Workshop on Applied Economics of Education) for useful suggestions. The authors have no conflicts of interest.

[*]University of Luxembourg, Luxembourg. Corresponding author. Email: chiara.costi@uni.lu.
[†]Lancaster University, UK; Magna Graecia University, Italy.
[‡]Universidad Autonoma de Madrid, Spain; Lancaster University, UK; IZA, Germany.

# 1 Introduction

It is well-established that ill-health at birth is systematically associated with a range of adverse outcomes across the whole life-cycle (Almond and Currie, 2011; Conti and Heckman, 2012; Almond et al., 2018). Low birth weight (often defined as weighting less than 2,500g at birth; OECD, 2019) is commonly used as a predictor of poor general health as it is strongly correlated with several short- and long-term negative health outcomes, including disability and mental ill-health (Currie and Hyson, 1999; Marmot, 1997; Figlio et al., 2014; Linsell et al., 2015; Mathewson et al., 2017); as well as broader detrimental outcomes such as lower educational attainment, increased probabilities of unemployment and lower lifetime earnings (Behrman and Rosenzweig, 2004; Case et al., 2005; Linnet et al., 2006; Black et al., 2007; Royer, 2009; Trickett et al., 2020). As a result, the identification of key factors affecting birth weight, especially those that are modifiable, is still among the most relevant policy issues in human development.

While there is a large literature on the determinants of birth weight, previous studies have rarely considered the potential role played by the intergenerational persistence of socioeconomic status on the birth weight of future generations. More specifically, though there is an extensive body of evidence suggesting strong intergenerational correlations in human capital (Currie and Moretti, 2003; Black et al., 2005; Doyle et al., 2005; Lindeboom et al., 2009), previous studies have not yet considered whether the intergenerational persistence in education levels across two generations would influence the birth weight of a third generation. Equally, despite known intergenerational associations in risky health-behaviours (Banderali et al., 2015; Hines et al., 2021; Osborne and Bailey, 2022), papers in this area traditionally focus on two generations and do not often consider potential effects on the health (or birth weight) of subsequent generations. Therefore, earlier empirical contributions on the effects of education and smoking behaviour on children's health have mostly addressed the link between parents (often mothers) and children. Yet, whether the intergenerational persistence of socioeconomic status across previous generations could be used to examine the initial health endowment of future generations has not been investigated empirically.

The main objective of this paper is to revisit the determinants of birth weight by exploiting intergenerational information across three generations. In particular, we employ intergenerational correlations in education and smoking behaviour between grandparents and parents to inform the effects of parental education and smoking on children's birth weight via a two-stage residual inclusion model (Terza et al., 2008; Wooldridge, 2014). This implies predicting the potentially endogenous parental education and smoking behaviour variables using grandparents' education and smoking, respectively, via first stage equations. We then use the first-stage residuals in corresponding second-stage equations to account for unobserved confounders while investigating the effects of parental education and smoking on birth weight. We estimate two-stage residual inclusion models separately for education and different parental smoking behaviours (maternal smoking during pregnancy and parental regular smoking) and a model including the effects of both parental education and smoking, as well as their interaction, into a single second-stage equation. The latter model allows exploring whether the effects of parental education and smoking behaviour on birth weight might influence each other. We explore the sensitivity of our results through several robustness checks, including the use of a sub-sample of deceased grandparents (i.e. excluding the presence of a direct effect of grandparents' smoking on children's development while in utero), and using grandparents' migraine as an instrument for parental education following the approach of Sabia and Rees (2011).

We draw rich multi-generational information from the National Longitudinal Study of Adolescent to Adult Health (Add Health). Add Health is a panel survey following a nationally representative sample of US high school students initially interviewed in grades 7-12 during 1994-95. This cohort of individuals is followed throughout adolescence and adulthood via four follow-up interviews (1996; 2001; 2008; 2018). It includes rich information on the main respondents' socioeconomic status, education, physical and psychological well-being, relevant behaviours as well as labour supply and family formation when they become adults. Importantly, Add Health encompasses three generations as, in addition to detailed longitudinal data on the main respondents (generation II), it also collects variables on the main respondents' parents (generation I) as well as their children (generation III).

We find strong intergenerational persistence in education and smoking across generations throughout all our first-stage equations. Results from our second-stage (main) equations show that higher parental education decreases the probability of low birth weight while parental smoking, especially mothers' smoking during pregnancy, increases it. Heterogeneity analyses further suggest that the positive effect of parental education might be mainly driven by fathers, particularly among white families. When the effects of parental education and smoking are jointly estimated, their combined effect does not appear to be statistically significant, potentially suggesting that the effects of education and smoking behaviours might be independent of each other. Results seem to be confirmed by models estimated on a subsample of main respondents whose parents were deceased before the birth of their children (therefore excluding a potential direct effect of grandparents' smoking on grandchildren's health) and using grandparents' migraine as an instrument for parental education.

This paper builds on and contribute to two strands of the literature, the ones on the determinants of birth weight and the intergenerational persistence of risky health-behaviours. First, we add to the existing literature on the factors affecting birth weight by exploring the effects of parental education and different smoking behaviours on low birth weight while accounting for the potential role played by the persistence across generations of socioeconomic status (via education) and risky health-behaviours (via smoking). While previous papers have estimated the effects of parental education and smoking behaviour, they mainly focused on two generations and have not used information on the intergenerational transmission of human capital or risky health-behaviours to aid the identification of parental effects. We also extend current analyses by including the effects of both education and smoking behaviours, typically considered separately, as well as their interaction. We explore such effects among same-sex descendants and by ethnicity (white versus non-white). Second, we contribute to the growing literature on the intergenerational transmission of risky health-behaviours by investigating whether the intergenerational persistence of smoking might affect the initial health capital of a third generation. Previous contributions on the transmission of smoking behaviour across generations rarely considered health effects on successive generations. In addition and differently from most earlier analyses, we go beyond the effects of mothers' smoking (during pregnancy) and also account for the effects of parental regular smoking, including fathers' smoking.

## 2 Literature

Our analysis bridges the literature on the determinants of birth weight, especially studies focused on the role of parental education, with the one on the intergenerational persistence of risky health-behaviours. Most previous studies of the first strand employ two generations of individuals and often emphasise the influence of mother's education on birth weight. These include the seminal paper of Currie and Moretti (2003) finding that higher maternal education (instrumented via the availability of colleges for women in US counties where mothers resided) increases children's birth weight, and subsequent contributions exploring the effects of parental socioeconomic status (as measured by income, education and occupation) on several child health outcomes, including birth weight (Case et al., 2002; Doyle et al., 2005; Currie, 2009; Lindeboom et al., 2009; McCrary and Royer, 2011). The overall message from this literature is that while higher parental socioeconomic status appears to generally improve child health outcomes and increase birth weight, estimates of the (causal) impact of higher parental education seems to be more mixed and may depend on the policies changes used to identify the effects of interest (e.g. the positive effect of Currie and Moretti, 2003, identified using policies promoting college entrances vs the null effects of Lindeboom et al., 2009, and McCrary and Royer, 2011, exploiting the 1947 raise of the school leaving age in the UK and variations in school entry policies in the US, respectively). In any case, these studies do not appear to have considered the persistence of socioeconomic status across multiple generations via education or risky health-behaviours and whether this could be exploited to aid the identification of the determinants of birth weight.

Although there is a vast and increasing literature on the intergenerational transmission of health (Ahlburg, 1998; Currie and Moretti, 2007; Coneus and Spiess, 2012; Johnston et al., 2013; Thompson, 2014; Maystadt and Migali, 2021; Bencsik et al., 2023), there are less studies on the persistence of risky health-behaviours across generations, including smoking and alcohol consumption (Göhlmann et al., 2010; Loureiro et al., 2010; Schmidt and Tauchmann, 2011; Brown and van der Pol, 2014). Though the latter studies tend to find strong intergenerational correlations between parents and children in smoking behaviour, they have not yet considered potential effects on the birth weight of subsequent generations.

Accordingly, we add to the literature by exploring the effects of parental education and different smoking behaviours on low birth weight. We do so by employing empirical models making use of the intergenerational persistence of education and smoking in previous generations for identification purposes. Our broader multi-generational approach may provide a more comprehensive view of the long-term, cascading effects of socioeconomic status and risky health-behaviors across generations. Furthermore, our analysis might contribute to the broader literature on the intergenerational transmission of inequality (Eshaghnia and Heckman, 2023) and dynastic human capital (e.g. Adermon et al., 2021), especially as the latter is focused on the influence of the generations prior to the parental one.[1]

# 3 Data

We employ multi-generational data from the National Longitudinal Study of Adolescent to Adult Health (Add Health). Add Health is a panel study representative of US students initially in high school and subsequently followed throughout adolescence and adulthood. Although Add Health currently includes five waves, given the objective of our analysis we focus on relevant data from the first four waves, that is when the main respondents are in grade 7-12 (wave I: 1994; wave II: 1995-1996) until they are aged 24-32 (wave III: 2001; wave IV: 2008).

Importantly, we exploit information across the first four waves to link each main respondent (generation II) with their parents (generation I) and their own children (generation III). As such, Add Health is uniquely suited to this multi-generational approach since it includes a wide range of variables on three generations. After matching all available information, we find that of the 20,745 main respondents initially sampled (generation II in the paper),[2] 8,234 individuals become parents of 17,137 children (generation III).

[1]The recent work of Eshaghnia and Heckman (2023) is relevant for this study as it suggests that maternal cognitive and social skills significantly influence educational choices, and these choices, along with maternal smoking habits, play a crucial role in determining the likelihood of delivering small-for-gestational-age babies (SGA). It also emphasises the relevance of maternal physical constitution as a strong determinant of delivering an SGA baby, particularly among lower educated individuals.

[2]Note that this initial sample of 20,745 students includes an initial core sample of 12,105 individuals plus extra samples of ethnic minorities; twins; individuals who were adopted; and other booster samples based on disability status as well as African-American students with highly educated parents.

## 3.1 Key variables

Add Health includes rich individual-level information covering three generations throughout two questionnaires: the in-home questionnaire and the parent questionnaire. While the former is a comprehensive survey answered in each wave by the main respondents, i.e. the parents (generation II) in our analysis, the latter was issued in wave I and collects sociodemographic information of the parents of the main respondents, i.e. the grandparents (generation I) in our case. Importantly, variables on grandparents from the parent questionnaire are complemented by additional questions answered by the main respondents as part of the in-home questionnaire in waves I and II. In addition, waves III and IV of the in-home questionnaire cover information about main respondents' biological children (generation III), including their birth weight. The birth weight of the parents (generations II) was collected as part of the parent questionnaire in wave I.

The main outcome of interest used in our models is low birth weight defined as a binary variable indicating whether children (generations III) were born weighing $\leq$ 2,500g (5lb 8oz) (OECD, 2019). As an alternative, birth weight measured as a continuous variable is also used in some of our robustness checks. All models also include children's biological sex as male newborns tend to weigh more and this is captured by a binary variable accounting for female children. While initially our models are estimated using any generations III children, further sensitivity tests include only first-born children as birth weight tend to increase with birth order (Bohn et al., 2021).

Parental (generation II) education is measured in years and based on the highest level of education achieved by each individual as reported by the main respondent (either the mother or the father) in the in-home questionnaire. Smoking during pregnancy is defined using information provided in waves III-IV and only refers to mothers' smoking behaviour.[3] The variable capturing regular smoking defines whether at least one of the respondents (either mothers or fathers) smoked a minimum of one cigarette on a daily basis in waves III-IV, before their children were born.

[3]The question about whether mothers smoked during their pregnancies is answered by main respondents (generation II) as part of the in-home questionnaire. When this question is answered by male respondents, mothers are the daughters-in-law of grandparents (generation I).

For the purpose of our analysis and to maximise the number of observations, information on grandparents' (generation I) education and smoking, is mainly drawn from waves I and II in-home questionnaires answered by the main respondents (generation II). The level of education of grandparents is defined using years of education corresponding to the highest education level of each grandparent. Following common practice in the education literature aimed at maximizing sample size (see Holmlund et al., 2011), our empirical models include a variable based on the sum of both grandparents' years of education. Grandparents' smoking status is defined as a binary variable, based on questions asking main respondents whether their mothers or fathers were regular smokers. This variable takes the value of 1 if at least one of the two grandparents was a regular smoker.

Our models also control for a wide range of observable characteristics specific to each generation. A binary variable for biological sex (female) is included as a covariate for children (generation III) together with a binary variable accounting for marital status (being married) for grandparents (generation I). In addition, as part of our robustness checks, we make use of a binary variable capturing whether at least one of the grandparents suffered from migraine headache. Parents' (generation II) covariates include the Peabody vocabulary test score from wave I (a standard test assessing general cognitive abilities, Dunn and Markwardt, 1970); three of the Big Five personality traits: extroversion, neuroticism and conscientiousness; and binary proxies for risky attitudes and myopic behaviour.[4] These sets of variables capture respondents' risk attitudes and noncognitive skills, which may also influence future choices around smoking and educational attainment. Other parental (generation II) relevant variables also include the biological sex of the main respondent, whether their own weight at birth was low ($\leq$ 2,500g); general health (a binary variable with value 1 for "excellent" or "very good health", 0 otherwise); and Body Mass Index (BMI), measured as a continuous variable on the BMI scale.

[4] Risky attitudes are defined via a binary indicator capturing at least one of the following behaviours: no use of seat belts or no use of birth controls. Myopic behaviour is measured using answers from a question asking main respondents whether they lived their life "without giving too much thought to the future", with answers on a 5-point scale. A binary variable is constructed, taking the value of 1 if respondents answered "strongly agree or agree", 0 otherwise.

### 3.2 Descriptive Statistics

Table 1 shows descriptive statistics of key variables for the three generations of individuals used in the analysis. Looking at generation I, while both grandparents present similar levels of education (around 14 years), a larger proportion of grandfathers appeared to have been regular smokers (61 vs 51 percent). Also, around 68 percent of grandparents were married and about 34 percent of them suffered from migraine headhache.

As expected, the average level of education of generation II, across both biological male and female individuals, as measured by the number of years spent in formal education, is slightly higher if compared to the one of generation I individuals. 14 percent of mothers smoke during pregnancy, whereas the percentage increases to 36 when looking at the variable capturing regular smoking among either mothers or fathers. 51 percent of individuals in generation II are white (vs non-white), although we should keep in mind that African-American adolescents (with highly educated parents) are oversampled (Harris, 2013). In addition, biological female individuals represent 60 percent of generation's II main respondents. Interestingly, 11 percent of main respondents were born with low birth weight.

Children (generation III) present an even proportion of males and females. Importantly, 10 percent of the children were born with low birth weight while the average birth weight is 7.19lb (corresponding to around 3,300g). This appears to be in line with standard birth weights measured in the US at the time (Currie and Moretti, 2007; Tilstra and Masters, 2020; Kennedy-Moulton et al., 2022).

## 4 Empirical Approach

Standard statistical associations between parental variables and children's birth weight might be biased by unobserved factors which may affect both. To improve on simple correlations and account for this potential issue, we exploit the intergenerational persistence in education and smoking between grandparents (generation I) and parents (generation II). We use this to aid the identification of the effects of parental education and smoking behaviour on children's (generations III) birth weight. More specifically, we follow Terza et al. (2008) and Wooldridge (2014) and employ a control function approach, based on a flexible two-stage residual inclusion model. This implies a first-stage equation estimating intergenerational correlations between grandparents' and parents' education and smoking, separately. This first-stage equation predicts the parental variable (either parental education or parental smoking) using among the predictors the corresponding grandparents' variable (grandparents' education or grandparents' smoking) and a series of other relevant controls. A second-stage equation includes the residuals from the first-stage equation and estimates the effects of the parental variable on children's birth weight. The main idea behind this approach is that the residuals from the first-stage equations account for the unobserved confounders affecting the parental variable of interest.

Although we employ two different sets of covariates in the first- and second-stage equations, since we estimate two linear equations our two-stage model is effectively equivalent to a standard two-stage least square estimator as well as to a two-stage predictor substitution model (Terza et al., 2008). The latter would imply the inclusion in the second-stage equations of the predicted values of the endogenous parental variables obtained during the same first-stage equations, rather than the corresponding residuals.[5] This type of control functions is based on an endogeneity test originally proposed by Hausman (1978).

The two-stage residual inclusion model implemented here takes the following form:

$$SES_{II} = \gamma_0 + \gamma_1 SES_I + \gamma_2 X + \epsilon \tag{1a}$$

$$LowBW_{III} = \beta_0 + \beta_1 SES_{II} + \beta_2 Z + \beta_3 \hat{\epsilon} + \xi \tag{1b}$$

Where $SES$ stands for socioeconomic status and indicates alternatively education or smoking behaviour of parents ($SES_{II}$) and grandparents ($SES_I$), respectively. The first-stage equation, Eq (1a), predicts $SES_{II}$ using $SES_I$ and $X$, which is a vector including variables related to generation I-grandparents (marital status) as well as a relevant set of covariates related to generation II-parents such as the Peabody vocabulary test, biological sex, risky attitudes, myopic behaviour, and three of the Big Five personality traits (extraversion, neuroticism and conscientiousness). In Eq (1b), the second-stage equation exploring the ef-

[5]Results obtained using two-stage least squares and two-stage predictor substitution models are available upon request as well as estimates produced using non-linear (probit) models for parental smoking behaviours in our two-stage residual inclusion models.

fects of the parental variables $SES_{II}$ (either education or one of the smoking behaviours) on children's birth weight $LowBW_{III}$, $Z$ is an additional set of variables related to generation II-parents that might be relevant controls for their children's birth weight, such as parental birth weight, general health and BMI. Moreover, the same equation also includes generation III-children's biological sex as this may play a role in influencing birth weight. Importantly, as outlined in Terza et al. (2008) and Wooldridge (2014), the second-stage main equation includes the residuals from the first-stage equation ($\hat{\epsilon}$) and the endogenous parental variable ($SES_{II}$).

We also consider, in a separate three-stage residual inclusion model, the combined effect of parental education and smoking on birth weight, including the effects of both parental education and smoking together with their interaction into a single second-stage equation. In this case, the two first-stage equations, for parental education and smoking, are estimated separately. That is, while equation (2a) estimates the intergenerational correlation between grandparents' education ($Ed_I$) and parental education ($Ed_{II}$), equation (2b) estimates the intergenerational correlation between grandparents' smoking ($Smok_I$) and parental smoking ($Smok_{II}$).[6] In the second stage, the combined effect of parental education and smoking is estimated on children's birth weight as equation (2c) shows, including residuals from both first-stage equations.

$$Ed_{II} = \gamma_0 + \gamma_1 Ed_I + \gamma_2 X + \epsilon \tag{2a}$$

$$Smok_{II} = \delta_0 + \delta_1 Smok_I + \delta_2 X + \mu \tag{2b}$$

$$LowBW_{III} = \beta_0 + \beta_1 Ed_{II} + \beta_2 Smok_{II} + \beta_3 Ed_{II} * Smok_{II} + \beta_4 Z + \beta_5 \hat{\epsilon} + \beta_6 \hat{\mu} + \xi \tag{2c}$$

Also in this framework, the endogenous parental variables ($Ed_{II}$, $Smok_{II}$) and their interaction are included in the second-stage equation (2c) together with the predicted residuals of the two first-stage equations ($\hat{\epsilon}$ and $\hat{\mu}$). Therefore, this approach allows us to estimate whether the effects of the two parental variables might interact or influence each other.

[6]Please note that in this model, education is defined as a binary variable taking the value 1 if respondents have completed at least 17 years of education (i.e. they are highly educated and completed college), 0 otherwise.

Effectively, we estimate five different residual inclusion models. Specifically, we estimate three separate two-stage residual inclusion models (of the type described in Eqs 1a-b), one looking at the effect of parental education on low birth weight and two investigating the effects of alternative smoking behaviours: mothers' smoking during pregnancy and parental regular smoking. In addition, we estimate two distinct three-stage residual inclusion models (as illustrated by Eqs 2a-c) exploring interactions between parental education and mothers' smoking during pregnancy versus parental education and parental regular smoking, respectively.

First-stage equations are estimated as either linear probability models (when the dependent variables are binary variables defining mothers' smoking during pregnancy or parental regular smoking) or standard linear models (when the dependant variable is parental education measured in years of education) whereas second-stage equations as generalized linear models using a maximum likelihood estimator (Wooldridge, 2010). Estimates of the second-stage equations are bootstrapped to compute reliable standard errors. The inclusion of the estimated residuals from the first stage in the second-stage equations provides a direct test of the endogeneity of our variables of interest (parental education and parental smoking behaviours), as discussed in Terza (2016, 2018). According to this test, the null hypothesis of exogeneity is rejected if the estimated coefficient of the corresponding residual is statistically significant.

# 5 Main results

Table 2 shows the results of the two-stage residual inclusion models considering the effects of parental education and smoking behaviours on children's birth weight separately. Panel 1 includes estimates from the first-stage equations while panel 2 the ones from the second-stage (main) equations. Each of the three models (looking at the effects of parental education, mothers' smoking during pregnancy and parental smoking, respectively) was estimated without as well as with covariates as reported in columns 1 and 2, correspondingly.

Estimates in columns 1 and 2 of Panel 1 show a strong intergenerational correlation between grandparents' (generation I) and parents' (generation II) education. An additional year of grandparents' education is positively correlated with parental education (increasing parental education between around 0.062 and 0.085 years), and the effect is highly statistically significant. In Panel 2, it can be observed that parental education (generation II) is negatively correlated with the variable defining children's (generation III) low birth weight. This implies that an extra year of parental education is associated with a reduced probability of children's low birth weight by between 1.1 and 1.3 percentage points (pp). Notably, the effect is larger and highly significant when including covariates (see column 2).

As displayed in Panel 1 columns 3-6, grandparents' regular smoking is associated with a higher probability (around 11pp) that mothers would smoke during pregnancy and that, more generally, parents would also regularly smoke (with an increase of around 18pp).[7] As a result, these findings also show the expected strong and highly statistically significant intergenerational effects. In Panel 2, we observe a positive effect of maternal smoking during pregnancy on the probability of low birth weight. This appears to suggest that maternal smoking during pregnancy is associated with an increased probability of low birth weight of around 10pp, and this effect becomes statistically significant at 1% after adding controls (see column 4). Estimates of the second-stage equations (Panel 2) for models focusing on parental regular smoking also present positive coefficients, though their effects are smaller and only weakly significant (see columns 5 and 6, respectively).

As for other covariates, parents' low birth weight (generation II) is statistically significant at 1 percent and presents a positive sign. This implies that there might be a certain degree of persistence in the intergenerational transmission of low birth weight (about 3.6pp), at least in part due to a shared genetic endowment, with potential implications on the perpetuation of low socioeconomic status across generations (Currie and Moretti, 2007). Interestingly, the parent's BMI does not appear to be statistically significant while the biological sex of the parent ("female") is associated with a slightly higher probability of children's low birth weight. Finally, parental (good) health only appears to be strongly statistically significant for the model looking at the effect of parental education.

[7] Note that since the question about mothers' smoking during pregnancy is answered by the main respondent (who could be either the father or the mother of the child), the effects of grandparents' regular smoking on mothers' smoking during pregnancy are estimated for mothers who could be either their daughters (if the main respondent was indeed the mother) or their daughters-in-law (if the main respondent was the father).

Looking at the residuals of the first-stage equations, as expected we reject the null hypothesis of exogeneity for parental education and maternal smoking during pregnancy, although corresponding estimates are highly statistically significant only in models with the full set of covariates. Residuals produced by the first-stage equation for regular smoking are not statistically significant and this may be also due to sample composition, since more than 50 percent of grandparents are regular smokers.

Table 3 presents results of the three-stage residual inclusion models exploring the effects of both parental education and smoking, their interaction and the full set of covariates. Here, parental education and smoking behaviours are binary variables indicating whether respondents are highly educated (i.e. have completed college/university) or smokers (i.e. either mothers' smoking during pregnancy or whether at least one of the parents is a regular smoker), respectively. In the first stage, the effects of grandparents' education and smoking behaviours (generation I) on corresponding parental variables (generation II) are separately estimated as first-stage equations. Looking at Panel 1, we still observe highly statistically significant and positive correlations between grandparents' education/smoking and parents' education/smoking behaviours.

Estimates in Panel 2 of Table 3 show that when parental education and maternal smoking during pregnancy are included in the same second-stage equation (column 1), both parental variables present highly statistically significant correlations with low birth weight. More specifically, highly educated parents are associated with a decrease in children's low birth weight of around 10.2pp, while mothers' smoking during pregnancy with an increase of about 10.6pp. Importantly, parental low birth weight is still highly statistically significant (3.6pp). Also, parental (good health) is correlated with a decreased probability of low birth weight (1.9pp). Overall, these findings appear to be line with the most recent evidence on the determinants of birth weight (Eshaghnia and Heckman, 2023). Column 2 includes estimates of the combined effects of parental education and regular smoking. In this case, only parental education appears to be statistically significant with a coefficient of similar magnitude to the one of column 1 (around 9.7pp). Notably, the interactions between parental education and the different smoking behaviours across the two three-stage models (columns 1 and 2 of Panel 2) do not seem to be statistically significant. This might indicate that while parental education and maternal smoking during pregnancy appear to affect birth weight, their corresponding effects might operate independently of each other. Moreover, the exogeneity tests based on the statistical significance of the residuals from the first-stage equations confirm previous results found in Table 2.

Table 4 shows results obtained replicating the baseline model presented in Table 2 by descendants of the same biological-sex (i.e. grandfathers-fathers-sons vs grandmothers-mothers-daughters). Here, Panel 1 explores intergenerational correlations between grandmothers and mothers (columns 1-3) as well as grandfathers and fathers (columns 4-5) for education and smoking behaviours.[8] We observe strong intergenerational correlations in education levels across descendants of both sexes, with estimated coefficients of between around 17.6-23.3pp (columns 1 and 4), which appear to be broadly in line with previous evidence on the intergenerational transmission of education (Holmlund et al., 2011). We also find strong and highly statistically significant intergenerational correlations of smoking behaviour across the first two generations regardless of biological sex (columns 2-3 and 4-5).

Panel 2 shows estimates of second-stage equations on the effects of parental education and smoking on low birth weight by grandsons and granddaughters separately. Results show that the effect of maternal (higher) education on female children's health does not appear to be statistically significant (Panel 2, column 1), while paternal (higher) education is associated with a highly statistically significant decrease in low birth weight by around 3.1pp (Panel 2, column 4). These findings point towards a positive influence of higher levels of paternal education on their children's birth weight. A potential explanation may be that higher paternal education increases household income through higher returns to education and assortative mating (and the latter appears to be confirmed by the similar levels of education achieved by both parents in our data, see Table 1). This may ultimately allow parents to access better health care or prenatal resources which might positively affect birth weight. In addition, it might also be possible that highly educated fathers could present a higher level of "health literacy" (i.e. the ability to understand and process health-related information) which has also been linked to better child health outcomes, (Jarosz and Gugushvili, 2020).

[8]Note that same-biological sex descendants models looking at male individuals only includes parental regular smoking, but not mothers' smoking behaviour.

Moreover, though we notice strong intergenerational correlation between mothers' and daughters' smoking behaviours in the first stage, in this case smoking during pregnancy does not appear to affect daughters' low birth weight. This might suggest that results found in the main analysis about the effects of smoking during pregnancy are mainly driven by male children. Also, the effect of parental regular smoking on birth weight appears to be predominantly influenced by fathers (Panel 2, column 5), although the corresponding coefficient is only weakly statistically significant.

Importantly, as evidenced by the recent study of Giuntella et al. (2023), both maternal and paternal low birth weight emerge as relevant factors affecting birth weight, although they are only strongly statistically significant (at 5 percent) when considering the effect of parental education (Panel 2, columns 1 and 4). Furthermore, maternal general health appears to be a consistently relevant control variable for low birth weight, implying that better maternal health is associated with a lower probability of low birth weight throughout all models looking at daughters (Panel 2, columns 1-3). It might be possible that the positive effect of overall good maternal health could partially compensate for the negative effects of smoking, thus explaining, at least to some extent, the non-statistically significant coefficients of maternal smoking during pregnancy and regular smoking. Indeed, it might be the case that such results are also driven by a relatively small sample size.

# 6 Robustness checks and heterogeneity analysis

This section proposes several robustness checks aimed at exploring the validity of our approach as well as the sensitivity of the main results. These include re-estimating our main models while minimising the potential direct effects of grandparents' smoking on children's health while in utero and the use of an instrumental variable for parental education. These additional findings are complemented by an heterogeneity analysis based on ethnicity (white versus non-white parents), further sensitivity tests on alternative sub-samples, and a discussion on the potential role played by the birth weight of parents included in all models as a relevant control.

## 6.1 Minimising fetal exposure to grandparents' smoking

One of the assumptions implied by our empirical approach is that grandparents' education and smoking behaviour should not have direct effects on their grandchildren's birth weight. In other words, grandparents' (generation I) education and smoking are allowed to influence their grandchildren's (generation III) health at birth only via the "middle generation", i.e. the parents of the children (generation II). While this might be reasonable to assume for grandparents' education, it might not be categorically excluded that some of the grandchildren's birth weights used in this study may be partly influenced by grandparents' smoking behaviour. For example, grandparents' smoking could directly affect grandchildren's low birth weight, if children were directly exposed to grandparents' smoking while in utero. For this reason and to minimise prenatal exposure to grandparents' smoking, the models for parental smoking behaviours are re-estimated by restricting the sample to those children whose grandparents passed away at the start of the study (see Table 5), that is before their grandchildren were born. This should avoid the direct influence of grandparents' smoking on their grandchildren's health.

Looking at Table 5, Panel 1 shows a strong intergenerational correlation between grandparents' (generation I) and parents' (generation II) smoking behaviour, both for smoking during pregnancy (column 1) and regular smoking (column 2). In Panel 2, results suggest that both maternal smoking during pregnancy and parental regular smoking are associated with increases in the probability of low birth weight by around 29pp and 20pp, respectively. Although the coefficients are less precisely estimated compared to the ones of the baseline models (and this might be due to the greatly reduced sample size), they appear to broadly confirm the main results.

## 6.2 Instrumental variable (IV) analysis for parental education

The validity of our results is further explored using an instrumental variable for parental education. Specifically, we exploit results from Sabia and Rees (2011) who estimated the effect of children's migraine headache on their educational outcomes using parents' migraine headache (specifically, maternal migraine headache) as an instrument on data from Add

Health.[9] Their approach implies that parents' migraine (the instrument) should affect children's education (the outcome) via children's migraine (the treatment). We leverage on the established correlation between parental migraine headache and children's education and use grandparents' migraine headache and as instrument for parental education in our analysis when estimating the effects of parental education on grandchildren's birth weight.

The results of this additional analysis are presented in Table 6. Panel 1 of Table 6 shows the first stage of this IV approach looking at the correlation between grandparents' migraine headache and parental education. As expected, grandparents' migraine headache is negatively correlated with parental education and this effect, according to Sabia and Rees (2011), should pass through parental migraine headache.[10] Results of the second stage of the IV approach are reported in Panel 2 and appear to suggest that parental education, instrumented with grandparents' migraine headache, contributes to a decrease of around 1.6pp of the probability of low birth weight. This also seems to confirm statistical significance and size of the effects found in our baseline estimates.

## 6.3 Heterogeneity analysis by ethnicity

Our baseline two-stage residual inclusion models are further re-estimated by focusing on the ethnicity of individuals in generations I and II, broadly defined as white vs non-white. Panel 1 of Table 7 shows that the intergenerational correlations of education and smoking behaviours between generations I and II are once again confirmed. Panel 2 shows results from the second stage equations suggesting that the effect of parental education on birth weight is only statistically significant for white parents (while for non-whites the corresponding estimate is still negative but not statistically significant). While this result appears to be interesting, we should point out that it could, at least in part, simply reflect the sample composition of Add Health, which includes a booster sample of African-American students with highly educated parents. As for smoking, this heterogeneity analysis seems to indicate

[9] Note that since their study was only focused on the relationship between parents and children, parents and children in their case would refer to generations I grandparents and generations II parents in our analysis, respectively.

[10] We also tested the correlation between grandparents' and parents' migraine headache directly on our data and found it to be strongly statistically significant. Estimates are available upon request.

that maternal smoking during pregnancy appears to be particularly detrimental for non-white children (around 30pp increase in low birth weight). This could be explained to some degree by the higher probability of maternal smoking during pregnancy that has been observed among non-white individuals (Yang et al., 2014).

## 6.4 Further tests and parental birth weight

We also performed additional tests to further ensure the overall reliability of our findings.[11] Since evidence from medical studies suggest that birth weight tend to increase with birth order (e.g. Bohn et al., 2021) we re-estimated our models by restricting our sample to generation III first-born children only, as in Currie and Moretti (2003). Results are very similar and do not appear to suggest a substantial penalty for first-born children in terms of a lower probability of low birth weight. We also replicated our main two-stage residual inclusion models using a continuous variable for birth weight as an alternative outcome. Once again, estimates reflect those obtained by our baseline models.

Finally, it should be noted that our models include the birth weight of the parents (more specifically, whether parents were also born with a low birth weight) as an important control. This is relevant as another assumption of our empirical approach exploiting multi-generational information is the absence of a direct effect of the grandparents' genetic endowment on their grandchildren's birth weight. Although specific genetic factors are strongly correlated with individual traits and health conditions, evidence suggests they often explain a small portion of the variation of children's health outcomes (Hirschhorn and Daly, 2005; Gibson, 2012; Price et al., 2015). In this case, grandparents' genetic traits related to education and smoking behaviour are not expected to have a strong residual direct effect on grandchildren's birth weight via their partly shared genetic endowment.[12] However, it might be possible that an interaction between genetic and environmental factors could have

[11] All corresponding results are available upon request.

[12] In the case of grandparents' education, this would imply that the level of education of a grandparent might influence their grandchildren's birth weight somehow directly and genetically and not just indirectly via the potentially improved socioeconomic status of their offspring (i.e. grandchildren's parents). In the case of smoking, this would assume that the genetic factors increasing the probability of smoking behaviour among the grandparents would have a residual direct effect on their grandchildren's birth weight not just through the smoking behaviour of their children (the grandchildren's parents).

a residual direct effect on birth weight (Jami et al., 2021; Bohacek and Mansuy, 2013). To account for the roles of such residual unobserved gene-environment interactions via grandparents, parental birth weight is included in all analyses. This should control for the genetic predisposition to low birth weight that, through the interaction with the overall environment, could directly affect our outcome of interest across generations. As already pointed out in the results' section, parental birth weight is always highly statistically significant and with a positive sign, implying that parents with low birth weight might have a higher probability of having children with low birth weight. As a result, the effect found for our key variables (parental education and smoking behaviours) should be net of these residual unobserved gene-environment interactions.

# 7 Conclusions

This paper explores the influence of parental education and different smoking behaviours on children's birth weight by exploiting rich intergenerational information across three generations from Add Health. Our main empirical approach is based on a two-stage residual inclusion model, which allows harnessing the intergenerational correlations of education and smoking behaviours to help identifying the effects of the parental variables on birth weight, while also accounting for unobserved confounders. We contribute directly to the current literature by providing a more comprehensive examination of different pathways through education and several smoking behaviours that may affect birth weight, a highly relevant early-life health outcome with major long-term effects, and exploiting multi-generational data.

In line with previous studies (e.g. Currie and Moretti, 2003), results suggest that higher levels of parental education may have a positive effect on children's health by reducing the probability of low birth weight and this is confirmed by an alternative model using an instrumental variable for parental education. Our heterogeneity analyses indicate that this appears to be driven mainly by the influence of fathers' education on their sons' birth weight and by white parents.

Also consistent with prior research (Pereira et al., 2022), maternal smoking during pregnancy seems to be associated with an increase in the occurrence of low birth weight. Nonetheless, this effect is only found to be highly statistically significant in our baseline models. Overall, parental regular smoking before and during the pregnancy appears to have smaller and only weakly statistically significant effects throughout all of our models. Importantly, a robustness check excluding a direct effect of grandparents' smoking on children while in utero, substantially confirm the larger correlation between mothers' smoking during pregnancy and low birth weight, yet estimates of both mothers' smoking during pregnancy and parental regular smoking are only weakly significant. A series of robustness checks such as using alternative sample restrictions and a different definition of the outcome variable also appear to support the main results.

We also provide estimates of three-stage residual inclusion models including effects of both parental education and, alternatively, one of the smoking behaviours (either maternal smoking during pregnancy or parental regular smoking) together with their interactions. Whereas these models confirm highly statistically significant estimates of higher parental education and mothers' smoking during pregnancy, there is no evidence of meaningful interactions between parental education and any of the parental smoking behaviours. This might suggest that education and smoking may influence low birth weight independently of each other. Importantly, all models also include parental low birth weight to account for the persistence of low birth weight across generations.

Our results might offer useful evidence to help informing policies aimed at improving child health outcomes and increasing birth weight. More specifically, interventions aimed at reducing school dropout coupled with health promotion initiatives targeting smoking prevention and cessation, especially during pregnancy, could increase awareness on the long-lasting impact that these choices can have on child health. While smoking cessation services should be especially designed for future mothers who are current smokers, initiatives to contain early exists from school could be particularly effective if targeted at male students, given that (future) fathers seem to play a major role in improving birth weight, especially through higher levels of education.

# References


Adermon, A., Lindahl, M., and Palme, M. (2021). Dynastic human capital, inequality, and intergenerational mobility. *American Economic Review*, 111(5):1523–1548.

Ahlburg, D. (1998). Intergenerational transmission of health. *The American Economic Review*, 88(2):265–270.

Almond, D. and Currie, J. (2011). Killing me softly: The fetal origins hypothesis. *Journal of economic perspectives*, 25(3):153–172.

Almond, D., Currie, J., and Duque, V. (2018). Childhood circumstances and adult outcomes: Act ii. *Journal of Economic Literature*, 56(4):1360–1446.

Banderali, G., Martelli, A., Landi, M., Moretti, F., Betti, F., Radaelli, G., Lassandro, C., and Verduci, E. (2015). Short and long term health effects of parental tobacco smoking during pregnancy and lactation: a descriptive review. *Journal of translational medicine*, 13(1):1–7.

Behrman, J. R. and Rosenzweig, M. R. (2004). Returns to birthweight. *Review of Economics and statistics*, 86(2):586–601.

Bencsik, P., Halliday, T. J., and Mazumder, B. (2023). The intergenerational transmission of mental and physical health in the united kingdom. *Journal of Health Economics*, 92:102805.

Black, S. E., Devereux, P. J., and Salvanes, K. G. (2005). Why the apple doesnât fall far: Understanding intergenerational transmission of human capital. *American economic review*, 95(1):437–449.

Black, S. E., Devereux, P. J., and Salvanes, K. G. (2007). From the cradle to the labor market? the effect of birth weight on adult outcomes. *The Quarterly Journal of Economics*, 122(1):409–439.

Bohacek, J. and Mansuy, I. M. (2013). Epigenetic inheritance of disease and disease risk. *Neuropsychopharmacology*, 38(1):220–236.

Bohn, C., Vogel, M., Poulain, T., Spielau, U., Hilbert, C., Kiess, W., and Körner, A. (2021). Birth weight increases with birth order despite decreasing maternal pregnancy weight gain. *Acta Paediatrica*, 110(4):1218–1224.

Brown, H. and van der Pol, M. (2014). The role of time preferences in the intergenerational transfer of smoking. *Health economics*, 23(12):1493–1501.

Case, A., Fertig, A., and Paxson, C. (2005). The lasting impact of childhood health and circumstance. *Journal of health economics*, 24(2):365–389.

Case, A., Lubotsky, D., and Paxson, C. (2002). Economic status and health in childhood: The origins of the gradient. *American Economic Review*, 92(5):1308–1334.

Coneus, K. and Spiess, C. K. (2012). The intergenerational transmission of health in early childhoodâevidence from the german socio-economic panel study. *Economics & Human Biology*, 10(1):89–97.

Conti, G. and Heckman, J. J. (2012). The economics of child well-being. Technical report, National Bureau of Economic Research.

Currie, J. (2009). Healthy, wealthy, and wise: Socioeconomic status, poor health in childhood, and human capital development. *Journal of economic literature*, 47(1):87–122.

Currie, J. and Hyson, R. (1999). Is the impact of health shocks cushioned by socioeconomic status? the case of low birthweight. *American economic review*, 89(2):245–250.

Currie, J. and Moretti, E. (2003). Mother's education and the intergenerational transmission of human capital: Evidence from college openings. *The Quarterly journal of economics*, 118(4):1495–1532.

Currie, J. and Moretti, E. (2007). Biology as destiny? short-and long-run determinants of intergenerational transmission of birth weight. *Journal of Labor economics*, 25(2):231–264.

Doyle, O., Harmon, C. P., and Walker, I. (2005). The impact of parental income and education on the health of their children.

Dunn, L. M. and Markwardt, F. C. (1970). *Peabody individual achievement test*. American Guidance Service, Incorporated.

Eshaghnia, S. and Heckman, J. J. (2023). Intergenerational transmission of inequality: Maternal endowments, investments, and birth outcomes. Technical report, National Bureau of Economic Research.

Figlio, D., Guryan, J., Karbownik, K., and Roth, J. (2014). The effects of poor neonatal health on children's cognitive development. *American Economic Review*, 104(12):3921–3955.

Gibson, G. (2012). Rare and common variants: twenty arguments. *Nature Reviews Genetics*, 13(2):135–145.

Giuntella, O., La Mattina, G., and Quintana-Domeque, C. (2023). Intergenerational transmission of health at birth: Fathers matter too! *Journal of Human Capital*, 17(2):000–000.

Göhlmann, S., Schmidt, C. M., and Tauchmann, H. (2010). Smoking initiation in germany: the role of intergenerational transmission. *Health economics*, 19(2):227–242.

Harris, K. M. (2013). The add health study: Design and accomplishments.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica: Journal of the econometric society*, pages 1251–1271.

Hines, L. A., Spry, E. A., Moreno-Betancur, M., Mohamad Husin, H., Becker, D., Middleton, M., Craig, J. M., Doyle, L. W., Olsson, C. A., and Patton, G. (2021). Cannabis and tobacco use prior to pregnancy and subsequent offspring birth outcomes: a 20-year intergenerational prospective cohort study. *Scientific Reports*, 11(1):16826.

Hirschhorn, J. N. and Daly, M. J. (2005). Genome-wide association studies for common diseases and complex traits. *Nature reviews genetics*, 6(2):95–108.

Holmlund, H., Lindahl, M., and Plug, E. (2011). The causal effect of parents' schooling on children's schooling: A comparison of estimation methods. *Journal of economic literature*, 49(3):615–651.

Jami, E. S., Hammerschlag, A. R., Bartels, M., and Middeldorp, C. M. (2021). Parental characteristics and offspring mental health and related outcomes: A systematic review of genetically informative literature. *Translational Psychiatry*, 11(1):197.

Jarosz, E. and Gugushvili, A. (2020). Parental education, health literacy and childrenâs adult body height. *Journal of biosocial science*, 52(5):696–718.

Johnston, D. W., Schurer, S., and Shields, M. A. (2013). Exploring the intergenerational persistence of mental health: Evidence from three generations. *Journal of health economics*, 32(6):1077–1089.

Kennedy-Moulton, K., Miller, S., Persson, P., Rossin-Slater, M., Wherry, L., and Aldana, G. (2022). Maternal and infant health inequality: New evidence from linked administrative data. Technical report, National Bureau of Economic Research.

Lindeboom, M., Llena-Nozal, A., and van Der Klaauw, B. (2009). Parental education and child health: Evidence from a schooling reform. *Journal of health Economics*, 28(1):109–131.

Linnet, K. M., Wisborg, K., Agerbo, E., Secher, N.-J., Thomsen, P. H., and Henriksen, T. B. (2006). Gestational age, birth weight, and the risk of hyperkinetic disorder. *Archives of disease in childhood*, 91(8):655–660.

Linsell, L., Malouf, R., Morris, J., Kurinczuk, J. J., and Marlow, N. (2015). Prognostic factors for poor cognitive development in children born very preterm or with very low birth weight: a systematic review. *JAMA pediatrics*, 169(12):1162–1172.

Loureiro, M. L., Sanz-de Galdeano, A., and Vuri, D. (2010). Smoking habits: like father, like son, like mother, like daughter? *Oxford Bulletin of Economics and Statistics*, 72(6):717–743.

Marmot, M. G. (1997). Early life and adult disorder: research themes. *British medical bulletin*, 53(1):3–9.

Mathewson, K. J., Chow, C. H., Dobson, K. G., Pope, E. I., Schmidt, L. A., and Van Lieshout, R. J. (2017). Mental health of extremely low birth weight survivors: A systematic review and meta-analysis. *Psychological bulletin*, 143(4):347.

Maystadt, J.-F. and Migali, G. (2021). The transmission of health across 7 generations in china, 1789–1906. *Journal of Health Economics*, 79:102493.

McCrary, J. and Royer, H. (2011). The effect of female education on fertility and infant health: evidence from school entry policies using exact date of birth. *American economic review*, 101(1):158–195.

OECD (2019). *Health at a glance 2019: OECD indicators*. Paris: OECD Publishing.

Osborne, J. B. and Bailey, B. A. (2022). Does it matter when i quit? could i just cut down some? links between trimester-specific smoking amount, preterm birth, and low birth weight. *Birth Defects Research*, 114(1):5–12.

Pereira, R. D., Rietveld, C. A., and van Kippersluis, H. (2022). The interplay between maternal smoking and genes in offspring birth weight. *Journal of Human Resources*.

Price, A. L., Spencer, C. C., and Donnelly, P. (2015). Progress and promise in understanding the genetic basis of common diseases. *Proceedings of the Royal Society B: Biological Sciences*, 282(1821):20151684.

Royer, H. (2009). Separated at girth: Us twin estimates of the effects of birth weight. *American Economic Journal: Applied Economics*, 1(1):49–85.

Sabia, J. J. and Rees, D. I. (2011). Individual heterogeneity and reverse causality in the relationship between migraine headache and educational attainment. *Economics of Education Review*, 30(5):913–923.

Schmidt, C. M. and Tauchmann, H. (2011). Heterogeneity in the intergenerational transmission of alcohol consumption: A quantile regression approach. *Journal of Health Economics*, 30(1):33–42.

Terza, J. V. (2016). Simpler standard errors for two-stage optimization estimators. *The Stata Journal*, 16(2):368–385.

Terza, J. V. (2018). Two-stage residual inclusion estimation in health services research and health economics. *Health services research*, 53(3):1890–1899.

Terza, J. V., Basu, A., and Rathouz, P. J. (2008). Two-stage residual inclusion estimation: addressing endogeneity in health econometric modeling. *Journal of health economics*, 27(3):531–543.

Thompson, O. (2014). Genetic mechanisms in the intergenerational transmission of health. *Journal of Health Economics*, 35:132–146.

Tilstra, A. M. and Masters, R. K. (2020). Worth the weight? recent trends in obstetric practices, gestational age, and birth weight in the united states. *Demography*, 57(1):99–121.

Trickett, J., Johnson, S., and Wolke, D. (2020). Behavioural and educational outcomes following extremely preterm birth: current controversies and future directions. *Emerging Topics and Controversies in Neonatology*, pages 367–385.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

Wooldridge, J. M. (2014). Quasi-maximum likelihood estimation and testing for nonlinear models with endogenous explanatory variables. *Journal of Econometrics*, 182(1):226–234.

Yang, T.-C., Shoff, C., Noah, A. J., Black, N., and Sparks, C. S. (2014). Racial segregation and maternal smoking during pregnancy: A multilevel analysis using the racial segregation interaction index. *Social Science & Medicine*, 107:26–36.

Table 1: Descriptive statistics

| | Mean | SD | Min | Max | N |
|---|---|---|---|---|---|
| **Generation I** | | | | | |
| Grandparents' education (sum of years) | 24.58 | 7.49 | 0 | 40 | 7465 |
| Grandmother education (years) | 13.84 | 2.69 | 0 | 20 | 6876 |
| Grandfather education (years) | 13.81 | 2.78 | 0 | 20 | 5882 |
| Grandparents' smoking (at least one) | 0.69 | 0.46 | 0 | 1 | 8065 |
| Grandmother smoking | 0.51 | 0.50 | 0 | 1 | 7816 |
| Grandfather smoking | 0.61 | 0.49 | 0 | 1 | 5898 |
| Married | 0.68 | 0.47 | 0 | 1 | 7011 |
| Grandparents' migraine (at least one) | 0.34 | 0.47 | 0 | 1 | 6776 |
| **Generation II** | | | | | |
| Parental education (years) | 15.01 | 2.15 | 9 | 23 | 7857 |
| Mother education (years) | 15.21 | 2.10 | 9 | 23 | 4734 |
| Father education (years) | 14.70 | 2.17 | 9 | 23 | 3122 |
| Maternal smoking during pregnancy | 0.14 | 0.34 | 0 | 1 | 8219 |
| Parental regular smoking (at least one) | 0.36 | 0.48 | 0 | 1 | 8192 |
| Vocabulary test | 63.28 | 10.25 | 0 | 87 | 7821 |
| Myopic behaviour | 0.14 | 0.35 | 0 | 1 | 8234 |
| Risky attitude | 0.40 | 0.49 | 0 | 1 | 8234 |
| Extroversion | 0.18 | 0.38 | 0 | 1 | 5476 |
| Neuroticism | 0.08 | 0.28 | 0 | 1 | 8219 |
| Conscientiousness | 0.15 | 0.36 | 0 | 1 | 8206 |
| Ethnicity (White) | 0.51 | 0.50 | 0 | 1 | 8234 |
| Female | 0.60 | 0.49 | 0 | 1 | 8233 |
| Low birth weight | 0.11 | 0.31 | 0 | 1 | 6586 |
| Overall good health | 0.91 | 0.29 | 0 | 1 | 8231 |
| BMI | 22.74 | 4.30 | 11.21 | 51.37 | 8045 |
| Migraine | 0.13 | 0.34 | 0 | 1 | 6932 |
| **Generation III** | | | | | |
| Low birth weight ( $\leq$ 2.5 kg) | 0.10 | 0.30 | 0 | 1 | 16707 |
| Birth weight (lb) (0.45-6.8 kg) | 7.19 | 1.38 | 1 | 15 | 16004 |
| Female | 0.50 | 0.50 | 0 | 1 | 16893 |

Table 2: The effects of parental education and smoking on low birth weight

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| ***Panel 1 - IGT generations I and II*** | | | | | | |
| | $YoEd_{genII}$ | | $Smokpreg_{genII}$ | | $RegSmok_{genII}$ | |
| $\text{YoEd}_{genI}$ | 0.085*** | 0.062*** | | | | |
| | (0.003) | (0.003) | | | | |
| $\text{RegSmok}_{genI}$ | | | 0.109*** | 0.106*** | 0.177*** | 0.179*** |
| | | | (0.006) | (0.006) | (0.009) | (0.009) |
| Controls | | ✓ | | ✓ | | ✓ |
| ***Panel 2 - Effects on Low birth weight*** | | | | | | |
| $\text{YoEd}_{genII}$ | -0.011** | -0.013*** | | | | |
| | (0.005) | (0.003) | | | | |
| $\text{Smokpreg}_{genII}$ | | | 0.089 | 0.103*** | | |
| | | | (0.054) | (0.033) | | |
| $\text{RegSmok}_{genII}$ | | | | | 0.056* | 0.044* |
| | | | | | (0.032) | (0.026) |
| Low $\text{BW}_{genII}$ | | 0.036*** | | 0.036*** | | 0.037*** |
| | | (0.010) | | (0.010) | | (0.010) |
| Gen $\text{Health}_{genII}$ | | -0.019** | | -0.017* | | -0.019* |
| | | (0.010) | | (0.010) | | (0.010) |
| $\text{BMI}_{genII}$ | | -0.001 | | -0.001 | | -0.001 |
| | | (0.001) | | (0.001) | | (0.001) |
| $\text{Female}_{genIII}$ | | 0.016*** | | 0.017*** | | 0.016*** |
| | | (0.006) | | (0.005) | | (0.005) |
| First-stage residual | 0.008* | 0.012*** | -0.083 | -0.102*** | -0.044 | -0.034 |
| | (0.005) | (0.003) | (0.055) | (0.034) | (0.032) | (0.026) |
| N | 11451 | 11451 | 12365 | 12365 | 12349 | 12349 |

Significance levels: *** 1% ** 5% * 10% bootstrap SE (reps 1000)

$\text{Controls}_{genI}$: married; $\text{Controls}_{genII}$: vocabulary; sex; extroverted; neurotic; conscientious; risk lover; myopic

Table 3: Combined effects of parental education and smoking

| | (1) | (2) |
|---|---|---|
| ***Panel 1 - IGT generations I and II*** | | |
| ***HighEd***$_{genII}$ | | |
| YoEd$_{genI}$ | 0.010*** | 0.010*** |
| | (0.001) | (0.001) |
| Controls | ✓ | ✓ |
| ***Smokpreg***$_{genII}$ | | |
| RegSmok$_{genI}$ | 0.104*** | |
| | (0.007) | |
| Controls | ✓ | |
| ***RegSmok***$_{genII}$ | | |
| RegSmok$_{genI}$ | | 0.181*** |
| | | (0.009) |
| Controls | | ✓ |
| ***Panel 2 - Effects on Low birth weight*** | | |
| HighEd$_{genII}$ | -0.102*** | -0.097*** |
| | (0.025) | (0.025) |
| Smokpreg$_{genII}$ | 0.106*** | |
| | (0.035) | |
| HighEd$_{genII}$#Smokpreg$_{genII}$ | -0.032 | |
| | (0.029) | |
| RegSmok$_{genII}$ | | 0.008 |
| | | (0.028) |
| HighEd$_{genII}$#RegSmok$_{genII}$ | | -0.010 |
| | | (0.021) |
| Low Birthweight$_{genII}$ | 0.036*** | 0.036*** |
| | (0.010) | (0.010) |
| General Health$_{genII}$ | -0.019* | -0.022** |
| | (0.011) | (0.011) |
| BMI$_{genII}$ | -0.001 | -0.001 |
| | (0.001) | (0.001) |
| Female$_{genIII}$ | 0.016*** | 0.016*** |
| | (0.006) | (0.006) |
| First-stage residual_Edu | 0.111*** | 0.106*** |
| | (0.027) | (0.025) |
| First-stage residual_Smokpreg | -0.108*** | |
| | (0.035) | |
| First-stage residual_RegSmok | | 0.001 |
| | | (0.029) |
| N | 11310 | 11292 |

Significance levels: *** 1% ** 5% * 10% bootstrap SE (reps 1000)
HighEd is binary indicating high education ($\geq$ 17 years of education)

Table 4: Heterogeneity analysis: same-sex descendants

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| ***Panel 1 - IGT generations I and II*** | | | | | |
| | ***Grandmothers - Mothers*** | | | ***Grandfathers - Fathers*** | |
| | $\text{YoEd}_{genII}$ | $\text{Smokpreg}_{genII}$ | $\text{RegSmok}_{genII}$ | $\text{YoEd}_{genII}$ | $\text{RegSmok}_{genII}$ |
| Mother's $\text{YoEd}_{genI}$ | 0.233*** | | | | |
| | (0.013) | | | | |
| Mother's $\text{RegSmok}_{genI}$ | | 0.155*** | 0.201*** | | |
| | | (0.012) | (0.014) | | |
| Father's $\text{YoEd}_{genI}$ | | | | 0.176*** | |
| | | | | (0.018) | |
| Father's $\text{RegSmok}_{genI}$ | | | | | 0.092*** |
| | | | | | (0.023) |
| Controls | ✓ | ✓ | ✓ | ✓ | ✓ |
| ***Panel 2 - Effects on Low birth weight of:*** | | | | | |
| | ***Daughters*** | | | ***Sons*** | |
| $\text{YoEd}_{genII}$ | 0.003 | | | -0.031*** | |
| | (0.006) | | | (0.008) | |
| $\text{Smokpreg}_{genII}$ | | 0.087 | | | |
| | | (0.063) | | | |
| $\text{RegSmok}_{genII}$ | | | 0.038 | | 0.244* |
| | | | (0.047) | | (0.125) |
| Low $\text{Birthweight}_{genII}$ | 0.033** | 0.028* | 0.029* | 0.055** | 0.025 |
| | (0.017) | (0.016) | (0.016) | (0.026) | (0.025) |
| General $\text{Health}_{genII}$ | -0.057*** | -0.041** | -0.044** | 0.005 | 0.010 |
| | (0.021) | (0.020) | (0.020) | (0.028) | (0.025) |
| $\text{BMI}_{genII}$ | 0.001 | 0.001 | 0.001 | -0.005*** | -0.004*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| First-stage residual | -0.005 | -0.082 | -0.043 | -0.677*** | -0.231* |
| | (0.006) | (0.064) | (0.048) | (0.008) | (0.125) |
| N | 5360 | 5360 | 5360 | 3168 | 3168 |

Significance levels: *** 1% ** 5% * 10% bootstrap SE (reps 1000)

$\text{Controls}_{genI}$: married; $\text{Controls}_{genII}$: vocabulary; sex; extroverted; neurotic; conscientious; risk lover; myopic

Table 5: Robustness check: deceased grandparents

| | (1) | (2) |
|---|---|---|
| ***Panel 1 - IGT generations I and II*** | | |
| | $\text{Smokpreg}_{genII}$ | $\text{RegSmok}_{genII}$ |
| $\text{RegSmok}_{genI}$ | 0.081*** | 0.196*** |
| | (0.021) | (0.026) |
| Controls | ✓ | ✓ |
| ***Panel 2 - Effects on Low birth weight*** | | |
| $\text{Smokpreg}_{genII}$ | 0.291* | |
| | (0.150) | |
| $\text{RegSmok}_{genII}$ | | 0.196* |
| | | (0.118) |
| Low $\text{Birthweight}_{genII}$ | 0.033 | 0.028 |
| | (0.032) | (0.032) |
| General $\text{Health}_{genII}$ | -0.060 | -0.060 |
| | (0.041) | (0.040) |
| $\text{BMI}_{genII}$ | -0.002 | -0.002 |
| | (0.002) | (0.002) |
| $\text{Female}_{genIII}$ | 0.014 | 0.019 |
| | (0.020) | (0.020) |
| First-stage residual | -0.304** | -0.163 |
| | (0.152) | (0.117) |
| N | 1442 | 1442 |

Significance levels: *** 1% ** 5% * 10% bootstrap SE (reps 1000)

Table 6: Robustness check: IV for parental education

| | (1) |
|---|---|
| ***Panel 1 - IGT generations I and II*** | |
| $\boldsymbol{YoEd}_{genII}$ | |
| Migraine$_{genI}$ | -0.336*** |
| | (0.040) |
| Controls | ✓ |
| ***Panel 2 - Effects on Low birth weight*** | |
| YoEd$_{genII}$ | -0.016*** |
| | (0.004) |
| Low Birthweight$_{genII}$ | 0.033*** |
| | (0.010) |
| General Health$_{genII}$ | -0.019* |
| | (0.010) |
| BMI$_{genII}$ | -0.001 |
| | (0.001) |
| Female$_{genIII}$ | 0.017*** |
| | (0.006) |
| N | 11383 |

Significance levels: *** 1% ** 5% * 10%

Table 7: Heterogeneity analysis: ethnicity

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| ***Panel 1 - IGT generations I and II*** | | | | | | |
| | | ***White*** | | | ***Non-White*** | |
| | YoEd | Smokpreg | RegSmok | YoEd | Smokpreg | RegSmok |
| $\text{YoEd}_{genI}$ | 0.070*** | | | 0.056*** | | |
| | (0.004) | | | (0.003) | | |
| $\text{RegSmok}_{genI}$ | | 0.125*** | 0.201*** | | 0.055*** | 0.126*** |
| | | (0.010) | (0.013) | | (0.007) | (0.011) |
| Controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| ***Panel 2 - Effects on Low birth weight*** | | | | | | |
| $\text{YoEd}_{genII}$ | -0.012*** | | | -0.008 | | |
| | (0.004) | | | (0.005) | | |
| $\text{Smokpreg}_{genII}$ | | 0.026 | | | 0.303*** | |
| | | (0.031) | | | (0.072) | |
| $\text{RegSmok}_{genII}$ | | | 0.060* | | | -0.010 |
| | | | (0.032) | | | (0.037) |
| Low $\text{Birthweight}_{genII}$ | 0.019 | 0.020 | 0.018 | 0.046*** | 0.046*** | 0.048*** |
| | (0.014) | (0.013) | (0.013) | (0.014) | (0.013) | (0.013) |
| General $\text{Health}_{genII}$ | -0.029** | -0.024* | -0.021 | -0.011 | -0.009 | -0.016 |
| | (0.014) | (0.013) | (0.013) | (0.014) | (0.014) | (0.014) |
| $\text{BMI}_{genII}$ | 0.001 | 0.001 | 0.001 | -0.003*** | -0.003*** | -0.003*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| $\text{Female}_{genIII}$ | 0.014** | 0.014* | 0.013* | 0.018** | 0.020** | 0.019** |
| | (0.007) | (0.007) | (0.007) | (0.009) | (0.008) | (0.008) |
| First-stage residual | 0.011*** | -0.019 | -0.036 | 0.006 | -0.295*** | 0.013 |
| | (0.004) | (0.032) | (0.032) | (0.006) | (0.074) | (0.039) |
| N | 8431 | 8431 | 8431 | 8706 | 8706 | 8706 |

Significance levels: *** 1% ** 5% * 10% bootstrap SE (reps 1000)

$\text{Controls}_{genI}$: married; $\text{Controls}_{genII}$: vocabulary; sex; extroverted; neurotic; conscientious; risk lover; myopic