ISSN 1503-299X

WORKING PAPER SERIES

No. 10/2009

PARENTAL INCOME AND RISK BEHAVIOUR AS DETERMINANTS OF ADOLESCENT HEALTH

Ivar Pettersen

Department of Economics

Norwegian University of Science and Technology
N-7491 Trondheim, Norway
www.svt.ntnu.no/iso/wp/wp.htm

PARENTAL INCOME AND RISK BEHAVIOUR AS DETERMINANTS OF ADOLESCENT HEALTH

By Ivar Pettersen, NTNU

Abstract

The relationship between health and income among adults is well established. Adolescent health and parental income has not received the same attention. In this study we look at household income both as a direct determinant of adolescent health and as an important factor in relation to behavioural patterns among adolescents. The data used is from two surveys conducted in a Norwegian county (n=9 000, which accounts for 90 % of the age-group 13-19 in a Norwegian county). The results indicate that income works partly through some health-promoting behaviours but, still has a significant direct effect on adolescent health. We also find that high household-income does not cushion the effect of health-deteriorating behaviour, but it strengthens the probability that adolescents take part in physical activity. Household income is important in terms of increasing the probability that adolescents actively participate in sports and physical activity in general.

1. Introduction

Health disparities in the general population are present. Within the adult population these differences are partly caused by exposure to unhealthy environments, genetics and lifestyle. Some of these determinants may seem unavoidable, while others are by choice. Most of the adolescent population has not been exposed to health-hazardous behaviours and contexts to the extent that the effects are irreversible. That makes them interesting in the sense that their actions are not necessarily a product of their own choices, but also a function of parental characteristics and home-environment. This study will look at the determinants of adolescent health with particular focus on parental income and the effects parental income has on the lifestyle-choices adolescents make. We will test if parental income has both a direct effect on the health of their offspring and an indirect effect on the impact of their lifestyle. Using a set of lifestyle-components we will estimate the importance of these behaviours on health and test if these effects are mediated by parental income. The hypothesis is that high parental income may proxy a healthier and more nurturing environment which acts as a defence against unhealthy lifestyles. In the process we estimate three effects. First, does parental income determine the adolescents' assessment of their health? Second, to what extent does parental income work to mediate the effects of lifestyle choices? Finally, in the process we also estimate the importance of parental income for the lifestyle-choices studied.

The data we use are from a survey called UNG-HUNT, which covers a county in Norway. The survey data has been merged with official data on annual income data for both parents. In the process we need to take into account that some of the relationships between adolescent health and lifestyles may be a reciprocal relationship. For instance, health and physical activity are strongly correlated, but it is difficult to establish a causal pathway from physical activity to health. Such issues are addressed using various empirical approaches to tentatively try to quantify the causal effect from lifestyle to health. The fact that parental income also may be endogenous with respect to the health of their offspring complicates the empirical analysis. We utilize IV-estimation, control function approaches and multivariate discrete analysis to address these empirical problems.

Previous research on the relationship between income and health has mostly been concerned with the relationship among adults (Pettersen, 2009). Socio-economic status (SES) have long been regarded an important determinant of health. The evidence of such relationships among adults is vast; see Smith (1999) and Wilkinson (2003) for reviews. The empirical evidence on adolescent health is much more limited. Studying the determinants for youth health is important for many reasons. First and foremost, it lessens one of the main challenges related to the relationship between income and health among adults, the possible reverse causality. Second, a lot of inherited diseases will not, when people are still young, have manifested themselves sufficiently to be an important cause of bad health, for instance cardio-vascular diseases. Third, evidence exist that health disparities in adolescence survive into adulthood (West and Sweeting, 2004). Case, et al. (2005) find that poor health in youth also is associated with lower earnings and investments in human capital in adult life. Learning about what affects the youth may provide some insight into relevant policies of counteraction. The focus of this paper is to both identify determinants for health among children / adolescents and test whether these determinants depend on household income.

An income-effect[1] for children and adolescents has been established in several other studies. In the influential work by Case, et al. (2002) parental income matter to children's health in the US. Goodman (1999) uses several self-reported measures on health among adolescents. Combined with information about parents' education, income and occupation, the evidence on a socio-economic gradient is not conclusive. Among six different proxies for health only self-reported health, obesity and depression show signs of being explained by parental income. The results hold when controls for socio-demographic factors are included. Currie, et al. (2004) use UK data to reach conclusion similar to Case et al., but find that the gradient in general is lower. An income-effect is also verified by Doyle, et al. (2007). They also highlight the importance of allowing household income to be endogenous to health. Utilizing instrumental variables-techniques they obtain much stronger effects of income compared to the baseline which assumed income to be exogenous.

---

[1] Often referred to as a gradient. The term gradient refers to a continuous effect of SES on health, and not a threshold effect.

Deaton and Paxson (1998) write: "There is a well-documented but poorly understood gradient linking socioeconomic status to a wide range of health outcomes". The potential channels are plenty. Linking health to consumption of health-services, indirectly identifying a path through which income works, may be relevant in the US, but not necessarily in Norway. The Norwegian health care system is universal and "free", but with some fixed fees associated with outpatient care and visits to the doctor. These fees do not apply to the examinations and treatments of young kids as all children below 14 years old retrieve free treatment within all levels of the health-care system. A study of Nordic adolescents by Reinhardt Pedersen and Madsen (2002) assessed the importance of parents' labour market participation in relation to children's health and wellbeing in a survey drawn from the five Nordic countries. Their results indicate that children with parents who have not been employed during the last six months have a higher probability of having recurrent psychosomatic symptoms. These results indicate that the Scandinavian model of welfare do not completely cushion against economic hardship. Deaton and Paxon's comment emphasize that the channel which income works through is unknown. Case, et al. (2002) considered income as a mediating effect on chronic conditions on self-assessed health. Their framework allowed for a decomposition of the effect of income either directly or indirectly through the severity of these conditions. They did find such effects on certain types of chronic illnesses. Parental characteristics are also identified as important in Call and Nonnemaker (1999). They look at adolescent self-rated health in relation to both individual and parental characteristics. Their findings suggest that parental education matters, along with their parent's assessment of their own health. Call and Nonnemaker's study includes a crude control for financial strain (binary control for financial problems) which fails to make significant impact. Propper and Burgess (2007) study income in relation to health with a sample of 0-7 yrs children. They do find that low income families tend to have children of poorer health, but this effect is found to work through other parental characteristics. This literature highlights parental characteristics as both direct and indirect determinants of adolescent health, even though the causality of such results always is open to discussion (as stated by most authors).

The evidence of health-deterioration from risk-behaviour, for instance, smoking is solid, both in terms of statistics and physiological effects. Within the epidemiology-literature lifestyle has not been regarded an important source of health-differences (Contoyannis and Jones, 2004). The traditional approach in this literature involves explaining health (or other outcomes) with some proxy for social class, controlling for age and gender. Social class is typically measures by income, education or work-category. Such parsimonious models are intuitively appealing, but are not capable of forming a platform for policy to reduce the variations in, for instance, morbidity or mortality. The relationship between health and socio-economic class captures a set of individual and contextual features which are both correlated with health and social class. In any given model-specification such unobserved heterogeneity will play a part. Contoyannis and Jones (2004) show how important it is to control for such heterogeneity to estimate the impact of lifestyle-factors. Adolescent risk-behaviours, or lifestyle-factors, have previously been studied from different angles. Clark and Loheac (2006) look at risk-behaviour explained by school peer groups. In some of their analysis they control for parental characteristic, including income. It turns out that higher parental income increases the probability of abusing alcohol. They also verify peer group effects in the consumption of alcohol. Aughinbaugh and Gittleman (2004) looked at risky behaviours as a function of maternal employment. Their hypothesis was that risk-behaviour in adolescence depend on parental time spent with their mother during the children's first 3 years. They found little or no evidence that maternal employment is important to risk behaviour.

What we focus upon in this paper are lifestyle-factors, both health-rewarding and health-deteriorating, that are allowed to be mediated or strengthened by household income. The lifestyle-factors used are a subset of those used in Contoyannis and Jones (2004)[2] and Balia and Jones (2008). The factors we do not include are stress, body mass index (BMI) and sleep-patterns. Stress and sleep-patterns are excluded because of lacking proxies, while BMI is excluded for reasons of potential endogeneity. Some of the lifestyle factors we do include also suffer from potential endogeneity-problems which we will deal with in the empirical analysis. This is especially relevant for our proxies of physical activity. We want to test if adolescents from rich household are more resilient to such behaviours than peers from poorer households. Income is allowed to have both direct effects on health and indirect effects through the probability of taking on certain behaviours. The framework used is based on Case, et al. (2002), which allows for decompositions of the part of the income gradient that works through behaviour into severity and prevalence effects. Does income represent a threshold into certain behaviours or are the effects of the behaviours dependent on household income-level? The behaviours we look at are nutrition (breakfast, dinner and lunch), physical activity (training weekly and participation in active sports), smoking, use of smokeless tobacco and alcohol-consumption. We feel that this study will contribute to a better understanding of how income works to improve health.

Theoretically the relationship between health and income can be illustrated using the Grossman model (Grossman, 1972). Grossman's model is concerned with the relationship between personal health and income. The behavioural foundation in the model could be extended to a consciousness about children's' health as well. Income works through two channels in Grossman's model. It is

[2] Referred to as the Alameda Seven, based on an epidemiological study in the county of Alameda, California 1965.

important both as a consumption good and an investment good. An individual invest in health to reduce the number of sick-days, which indirectly affect the level of consumption for the individual. Income works to increase the stock of health through the consumption of health-rewarding services[3]. Health is also a good in itself, meaning that whatever the individual does it will harvest more utility from it the healthier the individual is. For all the reasons adults would be concerned with investment into their own health they should be equally concerned about their children's health. The higher the children's stock of health is the fewer work-days will be lost due to sickness among the children. And, any leisure-activity, including other family members, will yield more utility the healthier the others are. Adults tend to bring their children along for leisure activities.

This paper is organized as follows. I start laying out the empirical approach, both with respect to the variables included in the analyses and the analytical framework. The data are described in the next section, including a brief analysis of health and some descriptive statistics. Empirical results follow, first using naive assumptions about endogeneity, followed-up by a more thorough treatment of the potential sources of bias. A brief summary ends the paper.

2. Analytical framework and empirical specification

Health is assumed to depend on a set of behavioural factors, some individual characteristics and family-characteristics including household income. We have income-data for both parents. In the data we are able to recognize different types of family structures. All respondents are part of a family which contains either both parents, on parent only or one parent and a stepparent. If one of the parents are absent from the household we really have no measure of family-income. This is relevant for about 10% of the sample.

Determinants of health

The baseline income-variable is H_INCOME, which is the sum of both parents income divided by the number of parents in the household, in other words, the average income.[4] We have no data on employment-status of the parents. Aughinbaugh and Gittleman (2004) looked at risky behaviours as a function of maternal employment. They find little or no evidence that maternal employment is important to risky behaviour. In our data we are not able to control for the number of hours worked,

[3] The original model defined this input as consumption of health care. According to Grossman (1999) this could might as well be interpreted as more general ways of improving health.

[4] Ideally we would like to correct household income for the number of children in the household. Unfortunately, the only information we have on the latter is a binary indicator for having more or less than two children.

so this study provides some justification of this omission. If the children / adolescents are supported by only one of their parents, we need to include additional controls for that. The variable SINGLE_PARENT indicates if the respondent lives in a single-parent household (*i.e.* no step-parents present). Income is measured in natural logs in order to allow for some non-linearity in the relationship between health and income.

Education is controlled for through one parent, that is, the parent with the highest educational level (EDUCATION) at the time of survey. This data is reported by the adolescent respondents in the second survey only (n=3100). The work by Lien, et al. (2001) suggest that adolescents not necessarily are able to provide accurate information about their parents' educational level, as their results indicate only a "fair" agreement between adolescents and parents with respect to parental education. The results from West, et al. (2001) are more promising in this respect. They estimate Kappa-statistics which indicate "good" or "very good" agreement between young and old. Further, the fact that the parent with the highest level of education does not necessarily live in the same household as the child, casts this variable into further doubt. Still, if we assume that both parents (normally) have regular contact with their children, the parents' education will probably have the opportunity to influence the children. Education is being measured along a five category ordinal scale, where the base-category represents primary school only, and the upper category represents university-education lasting 4 years or more. Reports on parental educational level are available for the second survey only.

Based on the questionnaire we have subjective knowledge of other aspects of family life and family behaviour. As part of the set of controls we use two indicator variables to control for family size (FAMILY_SIZE) and one variable to control for parental smoking habits (PAR_SMOKE). Another variable added to this set is PARENTS_DRUNK, which reports how frequently the adolescents have seen their parents drunk. The range is from never to weekly, spanning over 5 ordered categories. The purpose of this variable is to capture how conscious the parents are about their behaviour in regards to being the most important role model to their children.

Obvious individual characteristic are age (AGE) and gender (WOMAN). Age is controlled for through dummy-variables, ranging from 13 to 19 yrs in the unghunt1 sample and from 16 to 20 yrs in the second survey. Using dummies rather than a continuous measure of age provides a slightly better fit. Another variable (BGFRIEND) captures whether the respondent has a boy- or girlfriend. This captures both the mood of the respondent and also some aspect of maturity among the adolescents. In line with the same argument we include a binary control for respondents reporting not to live with their parents (LIVES_ALONE) and a proxy for social network capital (ENOUGH_FRIENDS).

Health-affecting behaviours

Lifestyle or behaviour, which are the terms used interchangeably in this study, is studied both as determinants of health and as an outcome determined by parental and individual characteristics. The latter analyses assume that individuals are faced with a choice of behaviours, partly based on differences in knowledge about the consequences of such behaviour. SMOKER is a binary variable that captures that the respondents smoke daily. We also control for respondents that used to smoke regularly, but who have managed to quit (USED_TO_SMOKE). The control for alcohol consumption is ALCOHOL. This is a binary control for respondents reporting to have been drunk 10 times or more. In Scandinavia the use of snuff (smokeless tobacco) is quite common, and increasing both among adolescents and adults. We control for use of snuff with the variable (SNUFF).

Among potential health-promoting activities we focus on physical exercise and nutrition. Two variables control for physical activity, TRAIN_WEEKLY is an indicator that the respondents are doing physical exercises as often as at least twice a week. In addition we have included a variable based on the respondents' reports on participation in organized sports (DO_SPORTS). Nutritional aspects are covered by the frequencies of eating breakfast, lunch at school and dinner on a regular basis (BREAKFAST_OFTEN, LUNCH_OFTEN and DINNER_OFTEN). A binary variable captures that respondents eat these meals more than 4 times a week.

Analytical framework

The primary purpose of this study is to identify the importance of parental income on adolescent health and behaviours. The hypothesis is that income may have both direct and indirect effects on health. First we model health as a function of individual characteristics, parental income and risk-factors. Second we recognize that adolescents may have different thresholds in relation to the probability of exposing themselves to risk-behaviour and also health-promoting activities. Behaviour is treated as a choice. Adolescents choose their lifestyle based on personal evaluations of utilities and preferences. This choice can be questioned as the contexts these adolescents live in exert pressure on their behaviour, but we do not address such issues. We test whether these choices and their impacts has anything to do with parental income and other characteristics. We follow the work of Case et al (2002) and adopt the following empirical specification for the latter study.

$$R = \alpha_0 + a_1 \ln y + X\delta^R + u^R \tag{1.1}$$

$$H = \beta_0 + \beta_1(\ln y - \overline{\ln y}) + \beta_2 R + \beta_3(\ln y - \overline{\ln y})R + X\eta^H + \varepsilon^H \tag{1.2}$$

Equation (1.1) expresses a direct effect of income on risks (R), where the term risk in general covers both behaviours which may be health-deteriorating and health-rewarding. Some "risk" are in fact positively associated with health. Equation (1.2) formalizes the effect of income on health. Income is also allowed to have interactive effects with risk. That is, the negative (or positive) effects from behaviours are allowed to vary with parental income-level. The rationale behind this is that high income households have attributes that can alter the effects of risk/anti-risk-behaviour. Even if income in itself may not be important, other characteristics which are correlated with income may be important. Income may capture other correlated attributes and have a direct effect. High income households have more options in terms of both where and how to live. This can make a difference in terms of the negative effects from alcohol, drugs and smoking. Allowing this kind of behaviour in a controlled, nurturing environment may offset some of the negative aspects of the risks. Or enhance the effect of health promoting behaviours. Income in equation (1.2) is measured as deviation from the mean income. This allows us to interpret the effect of risk-variables at mean income. Both equations control for family and individual characteristics.

3. Sample data and descriptive statistics

The data on children and adolescents are collected from two surveys named UngHUNT1 (UH1 hereafter) and UngHUNT2 (UH2).[5] The surveys are parallel surveys to the North-Trøndelag health study (HUNT), a comprehensive study collecting information on health and other characteristics among all adults in a Norwegian county. The first UngHUNT included all children and adolescents aged 13 to 19 years old. It consisted of a questionnaire, an interview and a medical examination which collected information on blood pressure, lung capacity, height, weight and hip-circumference. The questionnaire was completed during school hours, with 8984 respondents which constitute 90 percent of the relevant population. The age-group 13-19 include adolescents both in secondary-school and high school. The questionnaire was in general similar for both groups.

The second UngHUNT-survey was a follow-up study which was conducted in the same manner as the first. The main difference was in the invited respondents. The second survey invited all respondents in 7th or 8th grade in the first survey, a total of 4000. The respondents in both surveys may be identified and together these surveys represent a 2-period panel. In the analysis the data from the two surveys are pooled together. This means that some individuals are represented twice in the data. This is empirically treated as an inference-problem where potential within-individual dependence is controlled for.

Descriptive statistics

[5] Nord-Trøndelag Health Study (The HUNT Study) is a collaboration between HUNT Research Centre (Faculty of Medicine, Norwegian University of Science and Technology NTNU), Nord-Trøndelag County Council and The Norwegian Institute of Public Health.

The questionnaire contains questions about general health and more specific health problems. The main focus of this study is to use the traditional measure of self-assessed health (SAH). Table 1 shows the distribution of self-assessed health by age. Almost 90 % of the respondents assess their health as either good or very good. In the analysis below a narrow definition of good health will be used, defining "very good" as good health. An alternative definition, referred to as the "broad" definition of health, will define good health as health reported as either good or very good.

Table 1 Self-assessed health by age.

Age	Poor	Not good	Good	Very good
13	0.4 %	8.0 %	58.7 %	32.9 %
14	0.6 %	8.4 %	59.6 %	31.4 %
15	0.9 %	10.7 %	58.4 %	30.0 %
16	0.6 %	9.0 %	62.8 %	27.5 %
17	0.5 %	10.9 %	62.3 %	26.4 %
18	0.6 %	12.4 %	60.8 %	26.2 %
19	0.5 %	11.1 %	60.6 %	27.8 %
20	0.7 %	10.3 %	60.3 %	28.7 %
N	61	1075	6357	2999
Share of total	0.6 %	10.2 %	60.6 %	28.6 %

Self reported health is established as a strong predictor of adult health (Schou, et al. 2006). We do not know the predictive power of children and adolescents' assessments of health, but using the data we have enables us to perform reality-checks which involve comparing self-assessed health with other indicators of health problems. More specifically we consider self-assessed health in relation to self-reported chronic conditions. The presence of diagnosed chronic problems such as asthma, diabetes, epilepsy etc. are most likely known to the adolescents so we feel that this information represents valid tests of what we believe should affect a general assessment of health. Some of the assessments are less specific. For instance the responses in relation to impaired movement (H_movem) and body-pains (H_body) both are subject to individual reference-points. The results in Table 2 are from a regression of the probability of good health against a number of self-reported chronic conditions, including impaired vision and hearing. Both the narrow and broad definitions are used, and we separate samples by gender.

Most conditions have the expected negative effect on self-assessed health (SAH). The impact of impaired vision and hearing agree fairly well on SAH, across genders, especially using the broad definition of SAH. The existences of mental issues, bodily pains, impaired movement and asthma have similar effects on health across the six specifications. Allergies and epilepsy[6] does not seem to systematically affect the probability of self-reported good health, while the respondents with diabetes seem to assess their health above the average respondents. Epilepsy and diabetes are very rare conditions among adolescents so the estimates are driven by very few observations. Still, the overall results indicate that self-assessed health pass the reality-check and works as a composite health-outcome.

Table 2 Subjective specific health indicators as determinants of self-assessed health. Odds ratios from logit-regressions. Controlled for age and gender.

VARIABLES	Prev	Good health (narrow)			Good health (broad)		
H_movem	2.1 %	0.457***	0.317***	0.714	0.377***	0.289***	0.500***
H_body	6.5 %	0.288***	0.289***	0.290***	0.292***	0.324***	0.270***
H_mental	4.5 %	0.310***	0.389***	0.242***	0.276***	0.325***	0.255***
H_allergy	6.6 %	0.909	1.020	0.759*	0.942	1.158	0.787
H_asthma	12.3 %	0.647***	0.675***	0.609***	0.764***	0.699***	0.821
H_migrene	3.5 %	0.720**	0.624**	0.827	0.806	0.685	0.878
H_diabetes	0.5 %	1.879*	2.496**	1.106	4.410***	2.517*	na
H_epilepsy	1.0 %	0.826	0.883	0.673	0.623*	0.629	0.650
H_vision	12.2 %	0.743***	0.795**	0.696***	0.774***	0.795	0.765**
H_hear	1.2 %	0.546**	0.630	0.364*	0.374***	0.366***	0.337***
Sample		Full	Boys	Girls	Full	Boys	Girls
Observations		10492	5226	5266	10492	5226	5245

*** p<0.01, ** p<0.05, * p<0.1.

In Table 3 we report both the prevalence of behaviours and the effect on self-assessed health. The estimates reported are odds-ratios, controlled for age and gender. We observe that that the probability of reporting good health increases with physical exercise and better nutritional habits. Smoking, consumption of snuff and alcohol is associated with lower health-assessments. All estimates are significant at the 1 % level, and resemble gradients, rather than threshold-effects, most of the time. Exceptions are skipping school-lunch and smoking where we do not observe systematic decreases in probability of good health for increasing patterns of behaviour. The data we will use in the multivariate analysis will contain binary versions of the behavioural variables. Details of the definitions are covered above.

[6] Epilepsy may seem to be over-reported with a prevalence rate of 1 percent. Gjerstad and Taubøll (2003) report a prevalence of 0.7 % within the general population, but higher among the young and elderly.

Table 3 Frequencies of behaviour and effect of behaviour on health (narrow definition). Odds ratios from logit estimation controlled for gender and age.

Behaviour	Freq %	odds ratio	Behaviour	Freq %	odds ratio
Training			*Skips school-lunch*		
Every day	6	1.000	Every school day	2	1.000
4-6 days a week	21	0.695***	4-6 days a week	2	0.600**
2-3 days a week	38	0.349***	Do not bring lunch	10	0.460***
Once a week	16	0.174***	1-3 days a week	11	0.742*
Not every week	5	0.221***	Less freq or never	76	1.079
Not every 2 weeks	3	0.151***			
Not every month	5	0.128***	*Smoking*		
Never	5	0.169***	No	70	1.000
			No, but used to	11	0.486***
Participation in active sports			No, but used to daily	5	0.722***
Yes	49	1.000	Yes, sometimes	12	0.228***
Used to	38	0.383***	Yes, daily	2	0.385***
No	14	0.367***			
			Snuff (smokeless tobacco)		
Breakfast			Never	83	1.000
Every day	66	1.000	Yes, sometimes	9	0.723***
4-6 days a week	13	0.657***	Yes, but have quit	4	0.584***
1-3 days a week	11	0.520***	Yes, every day	4	0.650***
Less freq or never	10	0.464***			
			Been drunk		
Dinner			Never	24	1.000
Every day	74	1.000	Once	7	0.908
4-6 days a week	19	0.747***	Yes, 2-3 times	10	0.842*
1-3 days a week	7	0.451***	Yes, 4-10 times	13	0.807**
Less freq or never	1	0.248***	Yes, more than 10 times	46	0.616***

*** p<0.01, ** p<0.05, * p<0.1.

4. Empirical results

The importance of parental income is modelled in different steps. First we look at the direct impact on self-assessed health, allowing for various controls. This is done because we then can estimate the gross-income effect on health. The controls are stratified into individual characteristics, family characteristics, risk-factors and parental education. Parental education is not available for the entire sample. The second part of the estimation-procedure is based on equations (1.1) and (1.2) in the previous section. These equations express the effect of income both on health and through risk-behaviour. This initial analysis potentially suffers from endogeneity-problems, both related to income and certain behaviours. Extensions will address these issues using a set of different strategies to control for endogeneity.

Is household income important to health? Table 4 contains the estimated effect of income on health using different controls and sample. Columns (a) through (d) uses the full sample, but with different control-sets. Parental education is only available for the oldest respondents, and the estimation-results from this sample are reported in columns (e) through (g) in Table 4.

Household income is an important indicator for self-assessed health, and the effect is in general robust among the different samples. The estimation which stands out is the one in column (c). Adding controls for risk behaviour reduces the effect of income from 0.074 to 0.040. This result suggests that income captures some important aspects in relation to risk-behaviour, when risk-factors are not controlled for. Another result is the fact that added control for parental education does not alter the effect of household income, see models e) and f). In fact none of variables controlling for parental education are significant, casting doubt into the quality of this education-measure. That income has an effect, even when controlling for education, is not novel as Case, et al. (2002) also find this. There is a noteworthy difference in the estimated effect of log household income on health. Compared with Currie, et al. (2004) and Case, et al. (2002) the estimated effect of income is smaller[7]. Our highest estimate[8] is only half of what Currie, et al. (2004) found, indicating that household income is less important to health assessments in Norway compared to the UK.

The effects of the health-promoting variables are as expected. Weekly physical activity and regular, frequent meals are associated with increased health assessments. The presumed deteriorating effects are less convincing as using snuff and alcohol consumption does not seem to be important to health-assessments. Smoking, current or previous, have negative effects, current smoking, not surprisingly, more deteriorating than former. There are some differences between samples as health assessments among the older respondents in the second survey are not associated with frequent dinners. Apart from that result, the differences in point estimates are small, but the precision differs in favour of the full sample. Another interesting result is that the income-effect increase when we include family-characteristics in the model. This is an indication that the characteristics we use somehow are negatively correlated with income.

---

[7] Based on ordered probit estimates from Case, et al. (2002), Table 2 and Currie, et al. (2004), Table 1.

[8] Table **4**, columns (a) and (g).

Table 4 Self-assessed health (SAH) for adolescents and household income. Ordered probit estimates[9]

Dependent var:	Self-assessed health						
	UH1+UH2 pooled				Sample 2 only		
Variable	a	b	c	d	E	f	g
Household income	0.0941***	0.0736***	0.0395*	0.0546**	0.0742*	0.0669*	0.0983**
Train weekly			0.311***	0.320***	0.318***	0.317***	
Do sports			0.293***	0.286***	0.283***	0.275***	
Breakfast often			0.110***	0.120***	0.128**	0.122**	
Dinner often			0.0904***	0.0978***	0.0649	0.0662	
Lunch often			0.149***	0.169***	0.200***	0.197***	
Snuff			-0.0491	-0.0565	-0.0125	-0.00706	
Smoking			-0.399***	-0.414***	-0.328***	-0.328***	
Used to smoke			-0.204***	-0.207***	-0.149*	-0.152*	
Alcohol			-0.0413	-0.0532	-0.0507	-0.0522	
Individual charac	X	X	X	X	X	X	X
Family charac		X		X	X	X	X
Risk factors			X	X	X	X	
Parental education					X		X
Observations	7486	7311	7486	7311	3088	3088	3088

*** p<0.01, ** p<0.05, * p<0.1. T-statistics in parentheses, corrected for within respondent dependence. Individual-characteristics are age, a proxy for social capital, and an indicator for living alone. Family-characteristics are marital status, family-size, parental smoking and parental drinking habits. Respondents living with stepparents are excluded from the sample.

The previous research on adolescent health disagrees on the presence of an age-dependent income (or SES) gradient, but most of the empirical work seems to verify such an effect. In Table 4 the only information useful in this particular respect is the differences in income-effects between the full sample and second survey only. Even though the full sample includes adolescents from the entire age-span, the income-effect is significantly lower compared to the sample with older respondents only. Even if the purpose of the analyses is not to address this issue the results indicate that the effect of household income becomes more pronounced with age.

Health and risk-behaviour

Having established the relationship between income and health among children and adolescents, and identified a drop in the marginal effect of income after the inclusion of behavioural factors, a closer look at income's indirect effect is called for. We will now use the framework established above and estimate equations (1.1) and (1.2). I practise this means that we estimates two equations, both

[9] Using STATA 9.2

which are estimated using linear regression.[10] The outcomes of these equations are self-assessed health and behaviour, respectively. Income plays a part in both equations and risk-behaviour is a determinant for health. This approach represents a way of addressing income as a direct and indirect determinant of self-assessed health. Table 5 reports the estimates of household-income (column a) on several health-affecting behaviours together with the estimated effect of these different behaviours on self-assessed health (column b). Household income has a positive, significant effect on most types of behaviour with the exceptions of smoking, dinner- and breakfast-habits for which income seems unimportant. Surprisingly income has positive effects on both health-rewarding and health-deteriorating behaviour. Higher parental income increases the likelihood that children and adolescents use snuff and drink alcohol. This may be a reflection on the fact that acquiring alcohol costs money. Or, income may capture differences related to education. Statistics indicate that the alcohol-consumption in Norway is increasing in educational level. This legacy of highly educated parents may be captured by their children. Poorer households may also not be so lenient in relation to sponsoring such consumption, and that this factor may play an important role in determining that adolescent's consumption of such goods. The largest marginal effects are associated with physical exercise, in the model represented with TRAIN_WEEKLY and DO_SPORTS.

Table 5 The effect of risk factors on self-assessed health and the effect of income on risk factors. Health measure: self-assessed health, narrow definition. OLS-estimates

	(a)	(b)	(c)	(d)
Risk behaviour	Income-effect on risk-behaviour	Risk effect on health	Income-effect on health (ctrl risk)	Interaction income-risk
	α_1	β_2	β_1	β_3
Train weekly	0.0573***	0.185***	-0.00427	0.0409**
Do sports	0.0748***	0.180***	0.00365	0.0357*
Breakfast often	0.0153*	0.0872***	0.0400***	-0.0116
Dinner often	-0.00487	0.0517***	0.0114	0.0280
Lunch often	0.0158**	0.0887***	0.0124	0.0220
Snuff	0.0208***	-0.0632***	0.0311***	0.0254
Smoking	0.00657	-0.169***	0.0318***	0.00609
Used to smoke	0.00246	-0.0396**	0.0329***	-0.00939
Alcohol	0.0189**	-0.0630***	0.0293**	0.0107

***** p<0.01, ** p<0.05, * p<0.1** All regressions include controls for individual and family characteristics (see notes under Table 4 for details). Estimates are derived from the following equations:

$$R = \alpha_0 + a_1 \ln y + X\delta^R + u^R$$
$$H = \beta_0 + \beta_1(\ln y - \overline{\ln y}) + \beta_2 R + \beta_3(\ln y - \overline{\ln y})R + X\eta^H + \varepsilon^H$$

[10] OLS-estimates and marginal effects from probit-estimation did not differ significantly. Neither did the standard errors. For these reasons OLS-estimates were preferred even if the behavioral outcomes are binary.

All risk-factors are important to self-assessed health (β_2) and the qualitative effects are as expected. Smoking, frequent drinking and consumption of snuff is deteriorating to health, while physical exercise and good nutritional habits are promoting to health. Physical exercise and smoking are the most important determinants. At mean income, taking part in physical activities increases the probability of good health with 18-19 %. Smoking reduces the probability of being in good health by 17 %. The controls for risk-factors have a strong effect on the importance of household income. The estimated effect of income (β_1), conditional on risk-factors and ignoring the interaction-terms, varies from -0.004 to 0.04. The estimated effect, when controlling for weekly training is insignificant, and the corresponding effect is also small when controlling for participation in sports. It seems that participation in sports and regular training captures important features of the effect of parental income. Income has the strongest effect on these factors, and in turn it seems that income works *through* these factors to promote health. The interaction between income and risk-variables (β_3) reveals potential buffer-effects (or enhancing effects) from coming from a wealthy family (column d). Only two of these interactions are significantly different from zero, indicating that household income does not weaken or strengthen the impact of most of the factors studied. There may be a social gradient related to who exposes themselves to health-deteriorating behaviour, but when this kind of behaviour is present it makes no difference if your parents are rich or poor. The interaction-terms between parental income and physical activity are both positive and significant at conventional levels. This means that the effect of physical activity increases with higher household income.

Prevalence or severity?

The empirical framework we use allows us to decompose the income gradient on health into severity- and prevalence-effects. The intuition behind the decomposition is that adolescents from poor families may have a higher probability of health-deteriorating behaviour (prevalence) and/or that the behaviour may have a stronger impact on health (severity). Formally we let R represent a risk-behaviour, H is an indicator of good health and $\ln y$ represents the log of household income. Following Case, et al. (2002), the probability that a child is in good health can be expressed as:

$$(1.3) \qquad \begin{aligned} P(H \mid X) = {} & P(H \mid R=0, X)P(R=0 \mid X) \\ & + P(H \mid R=1, X)P(R=1 \mid X) \end{aligned}$$

Suppressing the X's the change in the probability of good health can by decomposed into the following three terms:

$$(1.4)\qquad \frac{\partial P(H)}{\partial \ln y} = \frac{\partial P(H \mid R=0)}{\partial \ln y} + \left[\frac{\partial P(H \mid R=1)}{\partial \ln y} - \frac{\partial P(H \mid R=0)}{\partial \ln y}\right] P(R=1) + \left[P(H \mid R=1) - P(H \mid R=0)\right]\frac{\partial P(R=1)}{\partial \ln y}$$

The first term is the effect of income on good health in the absence of risk-behaviour. The second term is the "severity effect", which is the additional effect of income on good health given risk behaviour. The third term, referred to as the "prevalence effect", captures that poor adolescents may have a greater chance of taking part in risk-behaviour. Equations (1.1) and (1.2) can be used to estimate the severity- and prevalence-effects. The severity effect is measured as $\beta_3 \overline{R}$, where $\overline{R}$ is the probability of risk-behaviour. Remember that β_3 represents the difference in income-effect between those exposing themselves to risk and those who do not. The prevalence-effect is defined by $\beta_2 \alpha_1$. Relating this to equation (1.4) we see that α_1 represents the income-effect on the probability of risk-behaviour, while β_2 represents the effect of risk-behaviour on health.

The regressions which form the basis for the decompositions into prevalence and severity are based on a linear probability model estimated using ordinary least squares. This approach represents a break with the traditional procedures for analysis of discrete data. Marginal effects from probit and ols-estimates are so close that we consider this approach valid. This is convenient as the interpretation of interaction-terms in non-linear analysis is complex, see Ai (2003).

Table 6 The decomposition of the health-gradient into severity and prevalence

	A	B	C	D	E
	Avg risk	Income effect[11]	Income effect (risk=1)	Severity	Prevalence
	$\overline{R}$	γ_1	β_1 / γ_1	$\beta_3 \overline{R} / \gamma_1$	$\beta_2 \alpha_1 / \gamma_1$
Train weekly	0.672	0.032	-0.13	0.86	0.33
Do sports	0.509	0.032	0.11	0.57	0.42
Breakfast often	0.799	0.032	1.25	-0.29	0.04
Dinner often	0.755	0.032	0.36	0.66	-0.01
Lunch often	0.875	0.032	0.39	0.60	0.04
Snuff	0.13	0.032	0.97	0.10	-0.04
Smoking	0.218	0.032	0.99	0.04	-0.03
Used to smoke	0.0673	0.032	1.03	-0.02	0.00
Alcohol	0.387	0.032	0.91	0.13	-0.04

[11] Since risks are not included in the regression, the effect of income on health does not vary.

Notes:

Column A: The share of adolescents with risk-behaviour ($\overline{R}$).

Column B: The effect of log-income (γ_1) on health based on a regression not controlling for risks. Based on:

$$H = \gamma_0 + \gamma_1 \ln y + X\rho^H + v^H$$

Column C: The income-effect on health when controlling for risk divided by column B: $\left(\frac{\beta_1}{\gamma_1}\right)$.

Column D: The severity-effect: defined as the share of adolescents with risk-behaviour multiplied by the effect of the interaction-term β_3, divided by column B $\left(\frac{\overline{R}\beta_3}{\gamma_1}\right)$.

Column E: The prevalence-effect: The effect of risk on health multiplied with the effect of income on risk, relative to column B $\left(\frac{\beta_2\alpha_1}{\gamma_1}\right)$.

Decompositions of the income-gradient is reported in Table 6. Column B in the table represents the income effect on health, not controlling for risk-behaviours. All other columns, except column A are calculated relative to this.

A value above 1 in column C indicates that income is even more important to health when the relevant behaviour is present and controlled for. This measure indicates that income works through risk-behaviour whenever the number deviates substantially from 1. For instance, including the effect of training weekly completely wipes out the effect of income on health in general (value=-0.13), while including smoking does nothing to change the effect of income (value=0.99).

The results identify weekly training and active participation in sports as the most important behaviours which income works through, but the income-gradient also works through nutritional habits. The other risks have direct impact on health to various degrees but they do not represent a channel which income works through. For physical activity the severity and prevalence effects explain the entire income-gradient, severity far more so than prevalence. The interaction term is part of this effect so the importance of this behaviour actually increases with household income. The effect of active participation in sports seems to be channelled both through prevalence and severity, close to equally divided between the two components. The severity effect captures that it to a larger extent allows rich children to participate in sports than children from poor families. Intuitively the severity channel makes more sense for participation in active sports rather than regular physical activity. Training can take many forms, some which are price rationed and some which are not. To actively participate in sports demands some amount of spending. Kids may need to be driven to soccer-practise, annual fees need to be paid and parents need to allocate time towards these activities. Regular meals, especially dinner and lunch, work through the severity-channel. This indicates that regular meals are more important for the poor than for the rich. These results are based on insignificant interaction-terms so the robustness of these results are in doubt.

5. Endogeneity and alternative approaches

The above analysis has been true, not to an economic tradition of analysis, but to an epidemiologic approach. This is obvious through the lack of addressing potential empirical problems caused by omitted variables, reverse causality, selection and unobserved heterogeneity. The exogeneity-assumptions all such analysis relies on may, to some extent, be violated. Even though one can argue that reverse causality from adolescent health to parental income may be of minor importance, the exogeneity of income still is in doubt. The question is, what does parental income capture? If we have omitted variables which are relevant to adolescent health, and correlated with parental income, we have biased estimates of the effect of parental income. One way of dealing with omitted variables is by utilizing instrumental variables, see Angrist and Krueger (2001). The same remedy can be applied to another obvious empirical challenge, namely the reverse causality between health and physical activity. The ability to train weekly and participate in sports will most probably depend on health.

To provide a result relevant to policy we need to be able to control for factors which are correlated with both health and income. An illustration of this is if both parents and offspring suffer from some (possibly the same) chronic disease[12]. If this disease decreases the chance that the adolescent is physically active, and if it has any impact on parental income, the presence of it may work as the third factor deriving the results. In practise the chronic condition is part of the residual, potentially correlated with parental income and adolescent physical behaviour. The relationship between the effects of these variables is crucial to the analysis, so it is important to eliminate some of the omitted variables bias which is potentially relevant to both. In fact, income may work as a proxy for a lot of unobserved factors, individual and contextual, which shape this relationship. The analysis above is unlikely to provide evidence of a causal relationship.

The endogeneity of physical activity poses an empirical challenge. Physical activity is a binary measure which is included both individually and as an interaction with household income. This means that we have to address non-linear relationships between endogenous variables. The following set of equations illustrates the empirical approach. Equation (1.5) is equal to the specifications used above (see equation (1.2)). Equation (1.6) is the same risk-equation used above

[12] Not necessarily genetically transferred diseases as they are very rare. The fact that parents and offspring live in the same environment and are exposed to the same contextual characteristics is probably more important in that respect.

with the addition of $Z^R\tau^R$, where Z^R is the instrument (peer-group behaviour). Income enters into the model both as deviations from the mean (log), and as log income. The endogeneity of both these definition are covered through a control function in equations (1.5) and (1.6).

$$(1.5)\ H = \beta_0 + \beta_1(\ln y - \overline{\ln y}) + \beta_2 R + \beta_3(\ln y - \overline{\ln y})R + X\eta^H + \varepsilon^H$$

$$(1.6)\ R = \alpha_0 + a_1 \ln y + X\delta^R + Z^R\tau^R + u^R$$

$$(1.7)\ (\ln y - \overline{\ln y}) = X\eta^y + Z^y\gamma^y + \varepsilon^y$$

We use three different approaches to deal with the potential endogeneity issues. Income and physical activity are assumed to be endogenous. The interaction between the variables complicates the model giving the endogenous variables a non-linear dimension. Below I will walk through three ways to address the issues. At the core of all lies the use of control functions together with other strategies to deal with endogenous regressors. Details on the different approaches are listed below:

a) Traditional IV and control functions. The potential endogeneity of income is handled using a control-function[13] while physical exercise enter through predictions from first step regressions. The interaction-term is not addressed[14]. Constructing this term based on the predicted values of its components is labelled forbidden regressions in the literature (Wooldridge (2002), page 250). Wooldridge and Imbens (2007) argue that control-function may be more robust whenever the endogenous variables enter in a non-linear fashion. In a linear model with one endogenous variable only the control-function approach yield the same point estimates as the traditional IV. The control function also serves as a traditional test of exogeneity of the regressor. The instruments used for physical activity are in fact the linear predictions from a probit-specification. The probit equation is identified by peer-group behaviour.
b) Control functions and selection. Endogeneity of income is controlled for using the residuals from a first step regression. Physical activity is addressed as a selection problem. The selection

[13] Alternative approach to 2- stage-least-squares. Based on the first stage estimation, calculate the residuals. These residuals then enter the 2nd step estimation as a regressor. It controls for the residual component correlated with the endogenous regressor, hence the name control function.
[14] Leaning on results from Fisher (1965), also referred to by Wooldridge (2001), identification of the equation without the non-linear terms imply identification of the equation with the non-linear term. Wooldridge (2001) p. 234 writes: "Models with interaction between exogenous variables and endogenous variables can be shown to be identified. Models with interactions among endogenous variables are also fairly easy to handle. Generally it is good practise to check whether the most general linear version of the model would be identified. If it is, then the nonlinear version of the model is probably identified."

equation uses peer-group behaviour to identify exogenous variation in physical activity. In practise control for selection bias is implemented by the inclusion of the inverse Mill's ratio IMR.

c) Bivariate probit. This specification allows for simultaneous estimation of health and physical activity. The system of equations are defined recursively, where health is determined by physical activity, but without health in the equation for physical activity. The applicability of this procedure is covered by Maddala (1983) p. 123 and Greene (2003), page 715. Wilde (2000) shows that such a system is identified even without exclusion restrictions but we do use peer-group behaviour to identify the effect of physical activity.

Predicting the probability of physical activity means we have to recognize that the estimates we retrieve of the effect of the behavioural variables are not average treatment effects, but rather local treatment effects (LATE). According to Imbens and Angrist (1994): "LATE is the average treatment effect for individuals whose treatment status is influenced by changing an exogenous regressor that satisfies an exclusion restriction". In other word, the estimated effect is relevant for the specific sub-sample for which the instruments help determine whether or not the treatment is taken.

In addition to the approaches listed in a) - c) we try other permutations of the strategies used. For instance, as an extension to a), we also try using a control function for physical activity, similar to Wooldridge and Imbens (2007), page 4. This latter approach does rely on a correctly specified probit-equation, which represent a more strict assumption than the IV-approach relies on. Under b) we add an interaction between the IMR and the control function for income. This serves as a control for selection correlated with unobserved heterogeneity.

There is no optimal validated approach to choose which is superior to others. Using IVs, control functions or controlling for selection represents trade-offs in terms of efficiency and robustness. According to Wooldridge and Imbens (2007) a systematic analysis which compares IV and control function approaches is lacking. We choose to report a number of specifications which combined may suggest a direction of the bias inflicted to our estimates when exogeneity of behaviour is assumed.

We do focus on physical activity only when we address potential endogeneity of behaviour. Other types of behaviour are assumed to be exogenous, but we realize that the validity of the other estimates rely on the assumption of exogenous income. If we establish that parental income is endogenous to adolescent health, the estimates presented in Table 5 and Table 6 may be biased.

Assuming that some of the behaviours are affected by the respondents' peers, we are able to instrument training weekly and participation in sports with the share of respondents doing so within the schools the respondents attend. Calculating the probability of the behaviour on a wider sample than the one used in the regressions minimizes the problem of mathematical constructs. The validity of such an instrument partly relies on the fact that for a given residential location pupils are not able to choose which school to go to. There is a one-to-one mapping of residence and school. Parents may choose residential area based on information on schools, but the evidence of such behaviour is limited. There was, when the surveys was conducted, no way the parents could assess the quality of schools, apart from hear-say and appearance[15]. A study by Fiva and Kirkebøen (2008) looked into capitalization on school performance (after performance indicators became publicly available). They were unable to establish any relationships between housing prices and school performance. The issue at hand is not about school performance, but potential problems related to the decision on where to live as a function of physical activity among peers. We suspect that information on such neighbourhood characteristics is hard to get by. Case and Katz (1991) found strong effects of social interaction in term of neighbourhood characteristics for alcohol and other behaviours. This results is in contrast to Evans, et al. (1992) whose results indicate that such effects disappear when the endogenous nature of choice of reference groups are taken into account. In our case the peer-group effects turned out to be strong predictors of behaviour and valid as instruments in terms of overidentification-properties.

Income is instrumented using house type, log municipality population, the share of municipal population with higher education and kindergarten coverage as instruments. The instruments which characterize regional size, educational level and labour-market vary only between the 24 municipalities within the county of North-Trøndelag. The house type of respondents may capture both the parents' income and age. The latter is unobserved and may be important to child health as mature parents may have the experience required to lead their hopefuls into the right path. House type could also be interpreted as a function of permanent income. Living in a villa may be a result of consistent high earnings. Either way these instruments need the scrutiny provided by tests of instrument validity.

In Table 7 we have reported the decomposition of the health-gradient into severity and prevalence through physical activity. The first set of results refer to regressions where both physical activity are treated as exogenous (as reported above), while the other sets of results are based on treating household income and physical activity as endogenous according to the specifications a)-c) above. For first-step diagnostics refer to Table A1 in the Appendix.

---

[15] No student-performance indicators where available to the public during or before the surveys. Now, this information is made available to the public.

Table 7 Severity and prevalence-effects when controlling for the potential endogeneity of physical activity and household income. All calculations based on marginal effects. Refer to Appendix for estimation-results.

		Income effect (risk=1)	Severity	Prevalence
	Exogenous behaviour and income			
	Trains weekly	-0.13	0.86	0.33
	Does sports weekly	0.11	0.57	0.42
a)	Endogenous behaviour (predicted) and income (CF)			
	Trains weekly	0.64	0.19	0.78
	Does sports weekly	0.63	0.19	1.07
b)	Selection into behaviour (IMR) and income (CF)			
	Trains weekly	0.22	0.30	0.40
	Does sports weekly	0.22	0.19	0.49
c)	Bivariate probit, endogenous income (CF)			
	Trains weekly	0.51	0.18	0.43
	Does sports weekly	0.63	0.08	0.36
a)+	Endogenous behaviour (CF) and income (CF)			
	Trains weekly	0.23	0.29	0.76
	Does sports weekly	0.22	0.19	1.01
b)+	Selection into behaviour (IMR) and income (CF) + IMR x CF			
	Trains weekly	0.23	0.27	0.40
	Does sports weekly	0.22	0.19	0.49

Refer to Notes below Table 6 for details on the decomposition.

The results in Table 7 show that our assumptions about the endogeneity of household income and physical activity are crucial to the magnitude and channel of effects. Treating both income and physical activity as exogenous we have already established that income seems to work through physical activity, and that the most important channel is severity. That is, physical activity gets more important with household income in terms of the impact it has on health evaluations.

This result changes when we control for the endogeneity of income and physical activity. We observe that that income still works through physical activity and that controlling for physical activity reduces the effect of income. Between 40 and 80 % of the income-gradient is explained by either training weekly or actively participating in sports. The severity-effect now is modest; indicating that the effect training has on health does not differ much between adolescents with poor or rich parents. The prevalence-effect, on the other hand, is much stronger. The interpretation of this is that household-income is a determinant of who takes part in such activities and who does not. The estimates based on exogeneity-assumptions indicated that the hurdle to participate was less important than severity.

Recognizing the endogeneity of income and physical activity, the decision to participate turns out to be about the only channel income works through. There is an economic threshold into health-promoting physical activity, but the effect of it makes does not depend on parental income.

The different approaches all point towards prevalence as the main source of effect, but the general importance of physical activity as a channel income works through does differ between the empirical specifications. Using combinations of control functions and controls for selection we explain almost 80 % of the income gradient. The bivariate-probit and the specification using predicted physical activity differs. Estimating both equations simultaneously, with the added bonus of controlling for unobserved heterogeneity; we get results which are qualitatively equal, but different in size. Physical activity still explains a substantial part of the income gradient but the amount has been reduced to 49 and 37 percent for training weekly and participation in sports respectively. The largest impact is still on prevalence, cementing that income represent a threshold into activities and not a requirement for gaining from them.

So what about the assumptions about exogenous income? In the initial analysis, reported in Table 5 and Table 6, we did base our estimates on exogenous income. If this assumption did not hold these estimates could be biased. The results in Table A2 and Table A3 in the Appendix do not suggest that income is endogenous to health, when we control for physical activity. The basis for this conclusion is that the control function term is insignificant. On the other hand we see that not controlling for physical activity renders income endogenous. That is, the control functions in column (2) in tables A2 and A3 are significant. Similar analysis (not reported) for the other types of behaviour differ in this respect. The analyses indicate that income is exogenous both to behaviours and to health. In short, this means that the results and decompositions reported in Table 5 and Table 6 stand. One could argue that we also should allow income to be exogenous to health also when we study physical activity. We have chosen not to do this as income needed to be treated as endogenous to physical activity. Realizing that these techniques also reduce bias do to measurement errors, correcting one estimate of income, while leaving the other untouched, did not seem appropriate. The fact that the income-effects do change substantially when we control for endogeneity suggests that income could be treated in a similar manner in both regressions.

6. Concluding remarks

Previous studies have identified household income as an important determinant for adolescent health. This study presents results which support this. We also try to shed some light on why income seems to be important. The framework utilized in this study opens for the possibility that household income has indirect effects through behaviour. The empirical evidence provided here suggests that this is the case only for indicators of physical activity. High-income household are more likely to have children participating in physical activity. The two measures of physical activity indicate that training weekly or active participation in sports is rationed through parental income, but the importance of such activities does not depend on household income. Our most conservative estimates indicate that 40 – 50 % of the income gradient may be explained through the prevalence of physical activity among the young (or factors correlated). For most other behaviours, household income level does not cushion or strengthen the effect of behaviours.

The relative importance of severity or prevalence depends on whether physical activity is treated as exogenous to health or not. Ignoring such issues reveals a strong influence through severity. That is, higher household income increases the yield of physical activity in terms of improved health-assessments. Controlling for the endogeneity of physical activity and income alters the channel income works through. Now the prevalence channel is most important, meaning that adolescents from high-income household are more likely to participate in sports. Participation in sports is both time-consuming and costly so the prevalence channel is intuitively more important than the severity-channel.

In the process of establishing these results we have estimated several relationships regarding adolescent health, behaviour and household income. Household income does have a direct effect on adolescent health. Even if our results indicate that physical activity is an important channel income works through, most estimates leave room for a direct effect. Household income also is important to adolescent behaviour. We do not claim that these results represent a causal effect, even though income for the most part can be assumed to be an exogenous variable. There seems to some features that are associated with income which both form the health-evaluations of the adolescents and the behavioural patterns.

References

Ai, C. 2003. "Interaction terms in logit and probit models." *Economics letters*, 80:1, pp. 123.

Angrist, Joshua D. and Alan B. Krueger. 2001. "Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments." *The Journal of Economic Perspectives*, 15:4, pp. 69-85.

Aughinbaugh, Alison and Maury Gittleman. 2004. "Maternal employment and adolescent risky behavior." *Journal of Health Economics*, 23:4, pp. 815-38.

Balia, Silvia and Andrew M. Jones. 2008. "Mortality, lifestyle and socio-economic status." *Journal of Health Economics*, 27:1, pp. 1-26.

Call, Kathleen Thiede and James Nonnemaker. 1999. "Socioeconomic Disparities in Adolescent Health: Contributing Factors." *Annals of the New York Academy of Sciences*, 896:1, pp. 352-55.

Case, Anne C. and L.F. Katz. 1991. "The company you keep: the effects of family and neighborhood on disanvataged youth." *NBER working paper 3705*.

Case, Anne, Angela Fertig, and Christina Paxson. 2005. "The lasting impact of childhood health and circumstance." *Journal of Health Economics*, 24:2, pp. 365-89.

Case, Anne, Darren Lubotsky, and Christina Paxson. 2002. "Economic Status and Health in Childhood: The Origins of the Gradient." *American Economic Review*, 92:5, pp. 1308-34.

Clark, Andrew E. and Youenn Loheac. 2006. ""It wasn't me, it was them!" Social influence in risky behavior by adolescents." *Journal of Health Economics*, In Press, Corrected Proof.

Contoyannis, Paul and Andrew M. Jones. 2004. "Socio-economic status, health and lifestyle." *Journal of Health Economics*, 23:5, pp. 965-95.

Currie, Alison, Michael A. Shields, and Stephen Wheatley Price. 2004. "Is the Child Health / Family Income Gradient Universal? Evidence from England." Institute for the Study of Labor (IZA).

Deaton, Angus S. and Christina H. Paxson. 1998. "Aging and Inequality in Income and Health." *The American Economic Review*, 88:2, pp. 248-53.

Doyle, Orla, Colm Harmon, and Ian Walker. 2007. "The Impact of Parental Income and Education on Child Health : Further Evidence for England." University of Warwick, Department of Economics.

Evans, William N., Wallace E. Oates, and Robert M. Schwab. 1992. "Measuring Peer Group Effects: A Study of Teenage Behavior." *The Journal of Political Economy*, 100:5, pp. 966-91.

Fisher, Franklin M. 1965. "Identifiability Criteria in Nonlinear Systems: A Further Note." *Econometrica*, 33:1, pp. 197-205.

Fiva, Jon and Lars J. Kirkebøen. 2008. "Does the housing market react to new information on school quality?" *Discussion Paper*: Statistics Norway.

Gjerstad, L and E Taubøll. 2003. "Hva er epilepsi?" *Tidsskr Nor Lægeforen*, 123, pp. 1359-61.

Goodman, E. E. 1999. "The role of socioeconomic status gradients in explaining differences in US adolescents' health." *American journal of public health*, 89:10, pp. 1522-8.

Greene, William H. 2003. *Econometric Analysis*: Prentice Hall.

Grossman, M. 1999. "The Human Capital Model of the Demand for Health." *NBER working paper series*.

Grossman, Michael. 1972. "On the Concept of Health Capital and the Demand for Health." *Journal of Political Economy*, 80:2, pp. 223-55.

Imbens, Guido W. and Joshua D. Angrist. 1994. "Identification and Estimation of Local Average Treatment Effects." *Econometrica*, 62:2, pp. 467-75.

Lien, N., C. Friestad, and K. I. Klepp. 2001. "Adolescents' proxy reports of parents' socioeconomic status: How valid are they?" *J Epidemiol Community Health*, 55:10, pp. 731-37.

Maddala, G. 1983. *Limited-Dependent and Qualitative Variables in Econometrics*: Cambridge University Press.

Pettersen, Ivar. 2009. "Health and Income in Norway. A search for a causal link." *Working paper*. Department of Economics, NTNU.

Propper, Carol and John Rigg Simon Burgess. 2007. "Child health: evidence on the roles of family income and maternal mental health from a UK birth cohort." *Health Economics*, 9999:9999, pp. n/a.

Reinhardt Pedersen, C. and M. Madsen. 2002. "Parents' labour market participation as a predictor of children's health and wellbeing: a comparative study in five Nordic countries." *J Epidemiol Community Health*, 56:11, pp. 861-67.

Schou, Morten Brix, Steinar Krokstad, and Steinar Westin. 2006. "Hva betyr selvopplevd helse for dødeligheten?" *Tidsskr Nor Lægeforen*:126, pp. 2644-7.

Smith, James P. 1999. "Healthy Bodies and Thick Wallets: The Dual Relation between Health and Economic Status." *Journal of Economic Perspectives*, 13:2, pp. 145-66.

West, Patrick and Helen Sweeting. 2004. "Evidence on equalisation in health in youth from the West of Scotland." *Social Science & Medicine*, 59:1, pp. 13-27.

West, Patrick, Helen Sweeting, and Ewen Speed. 2001. "We Really Do Know What You Do: A Comparison of Reports from 11 Year Olds and their Parents in Respect of Parental Economic Activity and Occupation." *Sociology*, 35:2, pp. 539-59.

Wilde, Joachim. 2000. "Identification of multiple equation probit models with endogenous dummy regressors." *Economics letters*, 69:3, pp. 309-12.

Wilkinson RG, Marmot M. 2003. "Social Determinants of Health: the Solid Facts." World Health Organisation.

Wooldridge, Jeffrey M. 2002. *Econometric Analysis of Cross Section and Panel Data*: The MIT Press.

Wooldridge, Jeffrey M. and G. W. Imbens. 2007. "Control Function and Related Methods." *What's New in Econometrics?* NBER.

Appendix

Table A1: First step diagnostics

First step diagnostics	
Income	
F-stat identification	41.16
P-val overidentification	0.3534
Train weekly	
F-stat identification	269.53
P-val overidentification	na*
Does sports weekly	
F-stat identification	324.76
P-val overidentification	na*

*Since we use one instrument only the model is exactly identified. Adding the other peer-group variable to the instrument sets, allowing for a test of overidentifying restrictions, yields p-values of 0.59 and 0.01 for train weekly and sports, respectively.

Table A2: Estimates which the calculations in Table 7 are based on. Trains weekly. OLS-estimates.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent:	Trains weekly				Self-assessed health				
Age 14	-0.0022	-0.0071	-0.020	-0.021	-0.015	-0.017	-0.016	-0.017	-0.013
Age 15	-0.049**	-0.058***	-0.030	-0.034	-0.013	-0.013	-0.013	-0.014	-0.0099
Age 16	-0.095***	-0.10***	-0.062***	-0.066***	-0.033	-0.032	-0.032	-0.032	-0.030
Age 17	-0.12***	-0.14***	-0.043**	-0.051**	-0.0044	-0.0039	-0.0033	-0.0045	0.0033
Age 18	-0.12***	-0.14***	-0.028	-0.038*	0.0077	0.0093	0.010	0.0088	0.016
Age 19	-0.17***	-0.19***	-0.0055	-0.012	0.053*	0.054*	0.055*	0.052*	0.059**
Age 20	-0.24***	-0.25***	-0.073	-0.079	0.0076	0.0088	0.0093	0.0069	0.014
Woman	-0.055***	-0.056***	-0.13***	-0.13***	-0.11***	-0.11***	-0.11***	-0.11***	-0.11***
Family size 2	0.017	0.018	0.016	0.016	0.0063	0.0083	0.0085	0.0086	0.0054
Family size 3	-0.013	0.0025	0.0043	0.011	0.0087	0.0090	0.0085	0.0091	0.0037
Single parent	-0.086***	-0.076***	-0.064***	-0.059***	-0.034*	-0.033*	-0.034*	-0.034*	-0.037**
Parents smoke	-0.056***	-0.044***	-0.056***	-0.052***	-0.033**	-0.035**	-0.034**	-0.035**	-0.036**
Seen parents drunk 2	0.019	0.017	-0.029**	-0.030**	-0.034***	-0.035***	-0.035***	-0.035***	-0.034***
Seen parents drunk 3	-0.00031	0.00047	-0.049***	-0.049***	-0.047***	-0.048***	-0.047***	-0.048***	-0.047***
Seen parents drunk 4	-0.0053	-0.0078	-0.048*	-0.048*	-0.039	-0.043*	-0.041	-0.043*	-0.039
Seen parents drunk 5	-0.02	-0.011	-0.020	-0.017	-0.0050	-0.011	-0.011	-0.011	-0.0078
Lives away from	-0.090***	-0.080***	-0.073**	-0.069**	-0.041	-0.039	-0.039	-0.040	-0.044
Has boy-/girlfriend	0.021	0.016	-0.026**	-0.028**	-0.030**	-0.032**	-0.032**	-0.032**	-0.029**
Social capital	0.061***	0.056***	0.081***	0.078***	0.055***	0.058***	0.058***	0.058***	0.057***
Constant	0.30***	-0.78**	0.37***	0.37***	0.12**	0.29***	0.29***	0.12**	0.11**
Income (ln) within			0.032***	0.090*	0.058	0.020	0.021	0.021	0.013
CF income		-0.15***		-0.060	-0.046	-0.026	0.022	-0.026	
Train weekly predicted					0.35***				0.36***
Interaction term					0.026	0.040**	0.036**	0.039**	0.026
Income (ln)	0.057***	0.20***							
Trains weekly						0.18***	0.18***	0.34***	
IMR						-0.11***	-0.11***		
Interaction IMR x CF							-0.076*		
CF train weekly								-0.098***	
Observations	7302	7311	7311	7302	7302	7302	7302	7302	7311
R-squared	0.033	0.032	0.037	0.038	0.042	0.073	0.074	0.073	0.042

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors.

Table A3: Estimates which the calculations in Table 7 are based on. Does sports weekly. OLS-estimates.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent:	Does sports weekly				Self-assessed health				
Age 14	-0.073***	-0.079***	-0.020	-0.021	0.0065	0.0038	0.0043	0.0055	0.0084
Age 15	-0.15***	-0.16***	-0.030	-0.034	0.016	0.015	0.016	0.018	0.020
Age 16	-0.22***	-0.23***	-0.062***	-0.066***	0.0076	0.0055	0.0060	0.0089	0.012
Age 17	-0.26***	-0.28***	-0.043**	-0.051**	0.044*	0.040	0.041*	0.044*	0.052**
Age 18	-0.29***	-0.32***	-0.028	-0.038*	0.072**	0.069**	0.070***	0.071**	0.081***
Age 19	-0.40***	-0.43***	-0.0055	-0.012	0.14***	0.14***	0.14***	0.14***	0.14***
Age 20	-0.41***	-0.43***	-0.073	-0.079	0.080	0.076	0.077	0.073	0.088
Woman	-0.037***	-0.039***	-0.13***	-0.13***	-0.11***	-0.12***	-0.12***	-0.12***	-0.11***
Family size 2	0.024*	0.024*	0.016	0.016	-0.0025	0.0030	0.0031	0.0031	-0.0035
Family size 3	-0.0070	0.012	0.0043	0.011	0.0038	0.0049	0.0046	0.0049	-0.0011
Single parent	-0.087***	-0.073***	-0.064***	-0.059***	-0.027	-0.030*	-0.030*	-0.031*	-0.030*
Parents smoke	-0.059***	-0.045***	-0.056***	-0.052***	-0.031**	-0.034**	-0.034**	-0.034**	-0.034**
Seen parents drunk 2	0.0032	-0.00054	-0.029**	-0.030**	-0.032**	-0.030**	-0.030**	-0.031**	-0.031**
Seen parents drunk 3	-0.016	-0.015	-0.049***	-0.049***	-0.043***	-0.043***	-0.043***	-0.043***	-0.044***
Seen parents drunk 4	-0.065**	-0.068**	-0.048*	-0.048*	-0.018	-0.020	-0.019	-0.021	-0.017
Seen parents drunk 5	0.0063	0.018	-0.020	-0.017	-0.0058	-0.015	-0.015	-0.014	-0.0084
Lives away from parents	-0.076**	-0.063*	-0.073**	-0.069**	-0.033	-0.037	-0.036	-0.040	-0.035
Has boy-/girlfriend	0.033**	0.028**	-0.026**	-0.028**	-0.037***	-0.038***	-0.038***	-0.037***	-0.036***
Social capital	0.075***	0.069***	0.081***	0.078***	0.048***	0.051***	0.051***	0.052***	0.049***
Constant	0.11	-1.25***	0.37***	0.37***	0.14***	0.32***	0.32***	0.14***	0.13***
Income (ln) within			0.032***	0.090*	0.057	0.020	0.021	0.020	0.015
CF income		-0.19***		-0.060	-0.043	-0.017	0.011	-0.017	
Sports weekly predicted					0.37***				0.38***
Interaction term					0.034*	0.034**	0.029	0.033*	0.036**
Income (ln)	0.075***	0.26***							
Does sports weekly						0.17***	0.17***	0.35***	
IMR						-0.11***	-0.11***		
Interaction IMR x CF income							-0.031		
CF sports weekly								-0.11***	
Observations	7311	7302	7311	7302	7302	7302	7302	7302	7311
R-squared	0.070	0.072	0.037	0.038	0.043	0.075	0.075	0.075	0.043

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors.