Do Measured and Unmeasured Family Factors Bias the Association Between Education and Self-Assessed Health?

Christiaan W. S. Monden

Abstract The association between educational attainment and self-assessed health is well established but the mechanisms that explain this association are not fully understood yet. It is likely that part of the association is spurious because (genetic and non-genetic) characteristics of a person's family of origin simultaneously affect one's educational attainment and one's adult health. In order to obtain an unbiased estimate of the association between education and health, we have to control for all relevant family factors. In practice, however, it is impossible to measure all relevant family factors. Sibling models are particularly appropriate in this case, because they control for the total impact of family factors, even if not all relevant aspects can be measured. I use data on siblings from a US study (MIDUS) and Dutch study (NKPS) to assess the total family impact on self-assessed health and, more importantly, to assess whether there is a family bias in the association between educational attainment and self-assessed health. The results suggest that there is a substantial family effect; about 20% of the variation in self-assessed health between siblings can be ascribed to (measured and unmeasured) family factors. Measured family factors, such as parental education and father's occupation, could account only for a small part of the family effect. Furthermore, the results imply that it is unlikely that there is substantial bias due to family effects in the association between education and self-assessed health. This strengthens the conclusions from prior studies on the association between education and self-assessed health.

Keywords Education · Self-assessed health · Siblings · Sibling models · Family

1 Introduction

One of the most persistent findings with regard to questions about social inequality and life chances is the educational gradient in health and mortality (Antonovsky 1967; Huisman

C. W. S. Monden (✉)
Department of Sociology, Tilburg University,
PO Box 90153, 5000LE Tilburg, The Netherlands
e-mail: c.w.s.monden@uvt.nl

et al. 2005; Lynch 2001; Von dem Knesebeck et al. 2006). People with more education report better health and have longer life expectancies than lower educated people. The educational gradient in health has been documented for a variety of societies and time periods (Von dem Knesebeck et al. 2006). These health differences are not only observed between the highest and lowest educated groups, but health differences are found between all consecutive steps on the educational ladder (Erikson 2001).

Despite the large number of studies on health inequalities, many fundamental questions remain concerning the nature and causes of the association between education and health (Banks et al. 2006; Marmot and Wilkinson 2005; Phelan et al. 2004). One of these questions concerns the role of the family of origin. On the one hand, there is a strand of literature that studies the association between the socio-economic position of the family of origin and adult health (Ben-Shlomo and Kuh 2002; Graham 2002; Van de Mheen et al. 1997). Studies in this literature suggest that the socio-economic position of the family of origin is important for adult health. We know, on the other hand, from sociological studies, that the family of origin also plays an important role in determining an individual's educational attainment. In this paper, these two lines of research are brought together by studying the total family impact on self-assessed health and, more importantly, by assessing to what extent the association between education and health is biased by measured and unmeasured family factors that affect both education and health. In doing so, two important issues are addressed that have received little attention in the literature so far.

First, only a few characteristics of the family of origin have been included in studies on adult health. Typically information on father's socio-economic position during the respondent's childhood was available (Hertzman et al. 2001; Power et al. 1998; Van de Mheen et al. 1997). Although this is an important characteristic of the family of origin, many other family factors might be important as well, such as, for example, genes, lifestyle, or the family's social network. Thus, the total effect of the family of origin on adult health might be much stronger than suggested by studies that can only incorporate father's social class. This lack of measured family factors raises two questions: How large is the total family impact on self-assessed health? (Or, in other words, to what extent do family factors explain differences in adult self-assessed health?) And: How well do measured characteristics of the family of origin cover the total effect of the family of origin on adult health?

The second issue that has received little attention so far, is the question to what extent the association between education and self-assessed health is biased if not all family factors are taken into account. A serious bias could exist if some of the family factors that determine adult health are also important for educational attainment. Part of the association between education and health could actually be the result of growing up in a particular family rather than a true effect of education on health.

Prior studies primarily aimed at quantifying the additional importance of childhood circumstances for adult self-assessed health (after controlling for the current situation). They did not address the question whether education and health have important family factors in common. For example, we know that parental social class strongly affects children's education attainment (Hout and Diprete 2006; Shavit and Blossfeld 1993) as well as their health status in adulthood (Power et al. 1998). As a result, not taking into account parental social class would overestimate the association between education and health. A bias could also result from family factors that are not measured in the sample. To advance our understanding of educational differences in health, it is important to know the extent to which the association between education and health reflects the 'unbiased effect' of schooling. This is especially important for explanatory studies, such as Lynch (2001),

Stronks et al. (1996) and Van Lenthe et al. (2002) that can not take all family factors into account.

One of the reasons that the above mentioned questions have received little attention is that the majority of studies analyses samples of unrelated individuals. Samples with such a design can not be used to answer questions about total family effects adequately. Instead, samples are needed that contain multiple persons per family (i.e., siblings or twins). If we observe two or more siblings per family, we can draw conclusions about the total family effect by studying the shared variation among the siblings' outcomes. Sibling analysis makes it possible to assess how much of the variation in health and education is linked to family factors and to estimate the relative importance of measured and unmeasured family factors. The difference between sibling data and twin data is that the latter, given a number of assumptions, allows splitting the total family effect into genetic and non-genetic family components. Sibling studies have been used in various fields, for instance to assess the total family effect and 'unbiased effects' of schooling for outcomes such as income (Griliches 1979), occupational status (Hauser and Mossel 1985), cultural consumption (Van Eijck 1996), social stratification and social orientations (Sieben 2001). So far, there has been little investigation of the total family effect for self-assessed health (but see the twin studies discussed below). The current study focuses on self-assessed health, which is the most frequently used health indicator in social sciences (Ferraro and Farmer 1999).

To sum up, I address the following questions: (1) How large is the total family effect on self-assessed health? (2) To what extent do measured aspects of the family's socio-economic status account for this total family effect? and (3) To what extent is the association between education and self-assessed health biased by measured and unmeasured family factors?

2 Theory and Prior Research

As the research questions of this study are explorative it does not seem informative to formulate precise hypotheses (e.g., a hypothesis on the exact size of the family effect on self-assessed health). However, building on insights from sociology and social epidemiology and prior research, I will formulate a general hypothesis for each of the three research questions.

Regarding the first research question, there is ample reason to assume that there is a substantial family effect on adult self-assessed health. One's health status at a particular moment in time is not simply the result of current circumstances and behaviour, but rather the result of exposure to various circumstances and behaviour over the life course (Lynch and Kaplan 1997; Wadsworth 1997, 1999). Many physical conditions and also behaviour early in life are shaped by the family of origin. Furthermore, there is probably a genetic impact of the family. Thus, we can distinguish three mechanisms to explain how family factors influence adult self-assessed health: genetic, material and cultural factors. Below each one is discussed briefly.

First, an obviously shared family factor is genetic endowment. If self-assessed health has important genetic determinants, then we would expect siblings to show considerable similarities in health. Twin studies, however, have yielded mixed evidence about the genetic family component of self-assessed health. Two American studies on male twins (Romeis et al. 2000, 2005), two Finish studies (Leinonen et al. 2005; Silventoinen et al. 2007), two Swedish studies (Harris et al. 1992; Lichtenstein et al. 1993), one Danish study (Christensen et al. 1999) and one Norwegian study (Roysamb et al. 2003) have estimated

to what extent the resemblance in self-assessed health between twins can be attributed to family factors (genetic factors and the shared environment) on the one hand and individual factors (non-shared environment) on the other. The importance of the family factors in these studies varies strongly by country, sex and age. The percentage of variance that can be explained by the genetic family factor ranges from 15% to well over 50% in the different samples. In one subsample no genetic component was found (Harris et al. 1992). Some studies find that non-genetic family factors are associated with self-assessed health, whereas others conclude that only the genetic family factor is important (Romeis et al. 2000, 2005). Overall, these studies support the idea that family factors partly determine self-assessed health, although it remains an open question to what extent this family effect can be attributed to genetic factors, material family factors or cultural family factors.

The second family factor concerns material circumstances. Several authors have argued that growing up in materially less advantageous circumstances can have lasting negative effects on health (Blane et al. 1996; Hertzman et al. 2001; Lundberg 1993; Power et al. 1998; Van de Mheen et al. 1997). Crowded and poor housing, poor nutrition, low access to (preventive) health care as well as financial stress could have long-lasting effects on health. These adverse material circumstances are strongly related to low socio-economic status. Longitudinal and cross-sectional studies have shown significant associations between adult health and childhood material circumstances as measured by father's socio-economic status.

Third, the family of origin may shape lifestyles and attitudes that are relevant for health. Important health related behaviours such as smoking, dietary habits and exercise are, in part, influenced by family culture in childhood and adolescence. Also, attitudes and beliefs about health and health care may be influence by the family of origin. In sum, the above-mentioned genetic, material and cultural factors related to the family of origin would lead us the *expect a low to moderate correlation between the health statuses of adult siblings*.

Next I turn to the second research question. If we find significant resemblance in self-assessed health and education among siblings, can 'standard measures' of family factors, such as father's socio-economic position, explain the whole family effect? In the literature, parental social class and parental education are the most frequently used family indicators (Ben-Shlomo and Kuh 2002; Power et al. 1998). In addition also indicators of financial hardship during childhood are sometimes used (Lundberg 1993). These family factors are also available in the data used for the current study. But apart from the measured family factors, many potentially important family factors remain unmeasured or are only partly picked up by the measures of social class. Relevant aspects of the family concerning housing, the neighbourhood, culture and lifestyles remain unobserved and are not (fully) captured by the standard socio-economic indicators. Again, it is impossible to formulate a hypothesis about exactly what proportion of the total family effect is captured by the measured factors. Therefore the following more explorative hypothesis is formulated: *measured socio-economic family characteristics explain a substantial part of the total family effect on self-assessed health*.

The final question is whether there is substantial overlap between family factors that determine adult health, on the one hand, and family factors that influence educational attainment on the other hand. If there is such overlap the association between education and health would be partly spurious. In other words, the association is biased by family factors in studies that do not or can not take all possible family factors into account. It seems reasonable to expect at least some overlap in the relevant family factors. The family's socio-economic position is likely to affect both outcomes, albeit possibly through different mechanisms. The health effects of the family's position may be more related to

the material circumstances whereas the effect on educational outcomes may be more related to cultural factors and ambition. Sociological research has shown a persistent association between parental educational level and father's occupation on children's educational attainment (Blau and Duncan 1967; Hout and Diprete 2006; Sieben 2001) and these factors are also related to adult health.

There might also be genetic factors that affect education and health simultaneously. There is empirical evidence suggesting that some genetic factors that are related to self-assessed health are also related to education (Lichtenstein et al. 1993). Lichtenstein et al. (1993) found a correlation of $r = 0.21$ between the genetic family factor for self-assessed health and that for education. They also found a correlation of $r = 0.26$ between the factor for shared family environment (i.e., non-genetic family aspects) related to self-assessed health and that related to education. This suggests that another part of the association between education and health is caused by common non-genetic family factors. Although these correlations are not high, they do suggest that the effect of education on health is partly spurious. Therefore, I will test the hypothesis that *the association between education and self-assessed health is upwardly biased if not controlled for measured and unmeasured family factors*.

To sum, the expectation is that there indeed is a family effect on self-assessed health, which is partly explained by parental education and occupation, and that the pure effect of education on health will be smaller, but still significant, after taking all possible family factors into account.

3 Data

There are only few data sets that include relevant information on education, health and family background for adult siblings. I use an American and Dutch data set.

3.1 US Sample

The US data come from the MacArthur Foundation Midlife Development in the United States survey (MIDUS) (Brim et al. 2004). This is a national telephone and mail survey carried out in 1995–1996 under the auspices of the John and Catherine T. MacArthur Foundation Network on Successful Midlife Development (ICPSR 2760). A sample ranging in age from 25 to 74 years, was recruited from a random-digit-dial sampling frame of the coterminous United States and oversampling in five metropolitan areas. One respondent was selected from each eligible household. The survey was carried out in two phases: a telephone interview followed by a self-administered mail questionnaire. The phase 1 response rate was 70%, and the conditional phase 2 response rate was 87%, with an overall response rate of 61%. There are 4,242 cases (primary respondents) in the main sample.

Non-twin siblings were enrolled by sending a postcard to all respondents, asking them to provide contact information for their siblings. Since the family study was a secondary aim of the project, no elaborate non-response procedures were employed. While the number of eligible respondents who provided the names and addresses of their siblings was low (20%), the cooperation rate for the sibling sample was high (69%). In total 951 siblings were interviewed. The number of siblings from a single family ranged from one to six, including one sibling from 272 families, two from 146 families, three from 75 families, four from 22 families, five from 10 families, and six from four families. I constructed all possible sibling-pairs and applied a family weight to give each family equal influence in

the analysis. Apparently sisters were more willing to give each other's addresses. This is in line with the finding that sister-sister relationships are stronger than brother-brother or mixed-sex sibling relationships (Lee et al. 1990). As a result there is an overrepresentation of females in the MIDUS sample.

MIDUS includes a standard single-item question about general health: "in general, would you say your physical health is… (1) excellent, (2) very good, (3) good, (4) fair or (5) poor?" Self-assessed health has repeatedly been shown to be a good indicator of general health (Ferraro and Farmer 1999; McHorney 2000). In a longitudinal population-based study Mossey and Shapiro (1982) showed that self-assessed health predicts mortality independently of age, sex, "objective health status", life satisfaction, income and residence. Moreover, the association of self-assessed health with mortality was stronger than the association between objective health status and mortality. Many other authors have reported similar results for a variety of populations, countries and age-groups (Benyamini and Idler 1999; Idler and Benyamini 1997).

Information provided by the primary respondent and siblings on their own educational attainment is used. Education is coded in years of completed education. The primary respondent also reported the educational attainment of both parents and the father's occupation at the time when the respondent was an adolescent ("Tell me about the job he had for the longest time during your adolescence—when you were 12–18 years old"). Father's occupation is converted to manual or non-manual. I use mother's education and father's occupation. Finally, I use the answer to a question about the family's financial situation: "When you were growing up, was your family better off or worse off financially than the average family was at that time? (1) A lot better off to (7) A lot worse off".

3.2 Dutch Sample

The data I use for the Netherlands come from a large-scale study of family relations: the Netherlands' Kinship Panel Study (Dykstra et al. 2005). The NKPS is a nationally representative survey among 8,155 primary respondents aged 18–79. Residents of care-institutions, penitentiaries, homes for the elderly, and holiday homes were excluded from the sample frame. The data were gathered in 2002–2004. Interviews were held with respondents at home using CAPI face-to-face interviews and self-administered questionnaires. The overall response rate was 45%, which is low compared to the US but about average for the Netherlands.

For each primary respondent who had living brothers or sisters, one sibling was randomly selected to be interviewed as well. The primary respondents were asked to provide the address and grant permission to contact the sibling. This permission was granted by 60% of the respondents who had a living sibling. The response rate among the siblings was 63%. I select siblings in the 24–75 age range to match the MIDUS sample. In total, information from 2,289 siblings is used.

The health question in NKPS is similar to the one in MIDUS except that the answer categories differ slightly: "(1) excellent, (2) good, (3) not bad and not good, (4) poor, or (5) very poor".

The primary respondent and sibling report their highest obtained diploma. This information is converted to total years of education. Additionally, the primary respondent was asked to report father's and mother's education and father's occupation when the respondent was about 15 years old. The occupational title is converted to manual or non-manual. NKPS also offers an indicator of parents' wealth: whether family was home owner or renting a house when the respondent was 15 years old. In the Netherlands, this is a good

Table 1 Descriptive statistics of the variables in the US sample (MIDUS) and Dutch sample (NKPS)

	US sample		Dutch sample	
	Mean/proportion	SD	Mean/proportion	SD
Age (years)	48.23	12.52	45.88	11.94
Male (0/1)	0.39		0.43	
Education (years)	14.49	2.69	12.46	2.94
Self-assessed health (1–5)	3.69	0.93	4.10	0.74
Father's education (years)	10.97	3.90	9.96	3.51
Mother's education (years)	11.76	2.99	8.91	2.72
Father manual occupational class (0/1)	0.49		0.48	
Advantageous material situation[a]	2.99	1.24	0.58	
N (individual respondents)	1,480		4,578	

[a] MIDUS range 1–7; NKPS dichotomous (0/1)

indicator of the family's wealth. Descriptive statistics for both samples are given in Table 1.

4 Methods

To assess the total family effect, Hauser-Mossel sibling-models (Hauser and Mossel 1985; Sieben 2001; Van Eijck 1996) are estimated. Originally, these models were designed to test the effect of family background on occupational attainment, but the technique can be used for any outcome.

The more important the family is, the higher the resemblance among brothers and sisters will be. Thus, sibling resemblance can be used as an estimate for the total effect of family factors (including all shared genetic, environmental, social, and cultural factors). The sibling model decomposes the variance in the dependent variable into a family component (also called 'between-family') and an individual component (also called 'within-family'). The first refers to the variance that siblings have in common, the second to individual deviations from the common variance. Structural equation models make it possible to assess family (between-family) and individual (within-family) regression coefficients, as depicted in Fig. 1. The square boxes in Fig. 1 indicate observed variables, the ovals indicate latent variables. The loadings of the latent variables are constrained to equal 1. The individual regression coefficient represents the individual association between education and health, controlled for all family effects. If there is no family effect, the family regression coefficient and individual regression coefficient will be equal. However, if the family regression coefficient is significantly larger, in a better fitting model, than the individual family regression coefficient, then this indicates a family bias in the effect of education on health. By adding measured family characteristics to the model, one can estimate how much of the total family-effect is captured by these characteristics. This way, hypotheses can be tested about the nature of, at least part of, the family's influence. The siblings models are estimated as structural equation models using Lisrel 8 (Jöreskog and Sörbom 1993).

Five regression models are estimated. The first two models (A1 and A2) are conventional regression models in which siblings' health is regressed on age, gender, education
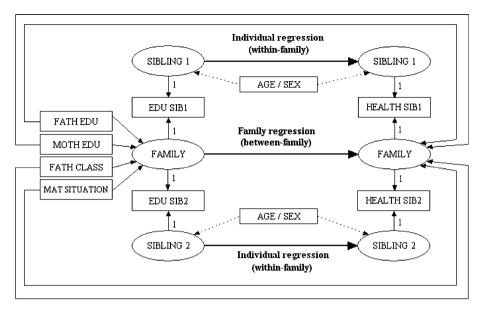
Fig. 1 Sibling model for the association between education and health. *Square boxes* indicate observed variables; *ovals* indicate latent variables in the structural equations model. The two individual regression coefficients (within-family regression) are set equal to each other

and four measured family characteristics. Because siblings are not independent observations, robust standard errors are used that correct for clustering within families (Rogers 1983; Williams 2000). These models are the reference models. The coefficients for education can be compared to the individual regression coefficient of education in the three siblings models (B1 to B3). The individual regression coefficient is set equal for both siblings in these models. The first sibling model, Model B1, assumes that the individual and family regression coefficients of education on health are equal. This equality constraint is relaxed in Model B2. If Model B2 has a significantly better fit than B1, there is a family bias in the educational effect on self-assessed health. In Model B3, I add observed family factors. How much the observed family factors contribute to the total family effect is indicated by the change (reduction) in the size of the family regression coefficient for education between Model B2 and Model B3. It is also reflected in the increase of the proportion of explained variance at the family level between the two models.

5 Results

5.1 Sibling Correlations

Table 2 shows the sibling correlations in education and self-assessed health. The results in this table are in line with earlier findings on sibling resemblance in education (Hauser and Mossel 1985; Lichtenstein et al. 1993). Four observations can be made from this table. First, sibling correlations for educational attainment are about twice as high as the sibling correlations for self-assessed health (the 95% confidence-intervals of the two correlation coefficients do not overlap). This suggests that the total family impact on adult self-assessed health is much smaller than the family effect on educational attainment.

Table 2 Age-adjusted sibling correlations of self-assessed health and education in the US sample (MIDUS) and Dutch sample (NKPS)

Type of sibling pair	US sample			Dutch sample		
	N	Health	Education	N	Health	Education
All	1,274	0.16	0.40	2,289	0.17	0.37
Mixed	619	0.11	0.32	1,054	0.14	0.31
Same sex	655	0.19	0.48	1,235	0.21	0.42
Brothers	256	0.11	0.42	450	0.24	0.33
Sisters	399	0.25	0.50	785	0.19	0.48

Second, sibling correlations tend to be higher among same-sex pairs than among mixed-sex pairs. This pattern holds for education as well as for self-assessed health. The correlation coefficients for education differ significantly between mixed sibling pairs and same-sex pairs in both samples. For health the same pattern is observed, but the correlation coefficients are not significantly different at the 5% significance level. Curiously, the sibling correlation for self-assessed health among US male sibling pairs seems much lower than the sibling correlation among female US pair and male and female Dutch sibling pairs. Note, however, that the difference between these correlation coefficients is only borderline significant (at a significance level of 10%).

The third observation is that sibling correlation for education is higher in female pairs than in male pairs. This difference is found in both samples and statistically significant in the Dutch sample. This suggests that the family is more important for women's education than for men's. Van Eijck (1996) showed that family background explains a larger proportion of variance in women's education compared to men in the Netherlands. However, this is most likely due to the smaller variance in women's education rather than to a larger impact of the family of origin (Van Eijck 1996).

Finally, we can observe in Table 2 that, despite some difference among the subgroups of sibling pairs, the overall correlations in the US and the Netherlands are quite similar; 0.16 and 0.17 for health and 0.40 and 0.37 for education in the US sample and Dutch sample, respectively. These small differences between the samples are not statistically significant.

In the remainder of this paper, I will analyse same-sex sibling pairs only. As mentioned above, the sibling correlations are systematically lower among mixed pairs compared to same-sex pairs (the differences are significant for education in both samples). The lower sibling correlations in mixed pairs suggests that the family is less important than the sibling correlations for male and female pairs imply. Thus, using the mixed sibling pairs raises important questions about differences between men and women in the variation in education and health and questions about how the family affects sons and daughters differently. As a result, mixed pairs would complicate the interpretation of the analysis. In order to obtain a clear picture about the impact of the family, I analyse male and female same-sex pairs simultaneously while taking into account age and sex. Splitting the analyses by sex would result in even smaller sample sizes. Working with small samples is not informative, because it increases the chance of type I or II errors. Studies on self-assessed health usually find relatively small effect sizes. Therefore, splitting the analyses into several subgroups would lead to imprecise estimates.

5.2 Conventional Models

Table 3 reports the outcome of conventional models (A1–A2) and siblings models (B1–B3) regressing self-assessed health on education for the US and Dutch sample. These analyses allow us to assess the total family effect on adult health and the 'unbiased' association between educational attainment and self-assessed health.

For the US sample, the conventional models show that there is a significant association between education and self-assessed health. Respondents with more years of schooling report better health than those with less education. The effect size is not particularly big. A 10 year difference in schooling corresponds to one standard deviation difference in health. The maximum educational difference is about 16 years; this difference in education corresponds to about 1.3 points on the 1–5 health scale.

By comparing the regression coefficients for education between Model A1 and Model A2, we can see whether the effect of education is overestimated in the conventional model if not corrected for observed family characteristics. Adding socio-economic characteristics of the family of origin reduces the coefficient for education by 20%, but this difference is not statistically significant. Given the relatively small sample size it is not surprising that the measured family factors do not reach statistical significance at the 5% level. The effects of mother's education and father's occupational class are both borderline significant.

5.3 Sibling Models

Next we turn to the sibling models. The first sibling model for the US sample constrains the individual and family regression coefficients for education to be equal. As can be observed from Table 3, the regression coefficient for education in Model B1 is almost equal to that in Model A1. The additional information that the sibling model gives us is a decomposition of the variance in self-assessed health. About one-fifth of the differences in self-assessed health among the respondents can be attributed to family factors. Note that this concerns all possible family factors; either social, genetic, observed or unobserved. Part of this family variance component is simply caused by the composition of families in terms of the siblings' age, sex and educational attainment. About 13% of the family variance can be ascribed to differences in composition. Most of the total variance consists of difference between individuals. About 9% of this remaining individual variance can be explained by age, sex and education.

In the second sibling model, the individual and family regression coefficients for education are allowed to vary freely. Relaxing the constraint does not significantly improve the model fit, as can been seen in the last two columns of Table 3. This suggests that, although the family coefficient for education is somewhat larger than the individual coefficient, there is no substantial family bias in the association between education and self-assessed health.

In Model B3, measured family factors are included in the sibling model. Adding these variables leads to a better fit compared to a model in which the effects of these variables are constrained to equal zero (chi-square of 81 with 32 degrees of freedom). The family coefficient for education is now much smaller and an additional 12% of the family variance in self-assessed health can be explained, mainly by mothers education and father's class (due to small sample size these effects do not reach statistical significance).

In addition to these three sibling models, I estimated Model B3 with an equality constrain on the individual and family regression coefficients for education. All results were similar to the unconstrained model except for the educational coefficient, which became

Table 3 Unstandardized regression coefficients (and standard errors) for same-sex sibling pairs from conventional OLS models and sibling models

	Regression coefficients		Age	Male	Father's education	Mother's education	Father manual	Material situation	Variance components				Fit measures	
	Education								Family level	Expl.[a]	Individual level	Expl.[b]	Chi2 (df)	GFI P-value
	Family	Individual												
US sample														
Model A1		0.082	−0.013	−0.030										
		(0.014)	(0.003)	(0.079)										
Model A2		0.065	−0.010	−0.051	0.003	0.025	−0.166	0.011						
		(0.014)	(0.003)	(0.077)	(0.013)	(0.015)	(0.090)	(0.029)						
Model B1		0.080	−0.013	−0.029					0.174	0.156	0.718	0.652	10.4	0.992
		(0.014)	(0.003)	(0.082)					20%	13%	80%	9%	(16)	0.844
Model B2	0.074	0.091	−0.013	−0.021					0.179	0.092	0.713	0.431	10.4	0.992
	(0.023)	(0.035)	(0.003)	(0.085)					20%	13%	80%	9%	(15)	0.797
Model B3	0.073	0.048	−0.010	−0.058	0.006	0.027	−0.170	0.015	0.186	0.088	0.689	0.431	13.3	0.993
	(0.023)	(0.047)	(0.003)	(0.083)	(0.015)	(0.017)	(0.090)	(0.035)	21%	25%	79%	6%	(23)	0.945
Dutch sample														
Model A1		0.022	−0.010	0.058										
		(0.005)	(0.001)	(0.033)										
Model A2		0.021	−0.009	0.061	0.007	−0.004	0.043	0.046						
		(0.005)	(0.001)	(0.033)	(0.006)	(0.007)	(0.037)	(0.033)						
Model B1		0.021	−0.100	0.058					0.094	0.093	0.436	0.416	20.1	0.996
		(0.006)	(0.001)	(0.033)					18%	1%	82%	5%	(16)	0.215
Model B2	0.022	0.020	−0.100	0.058					0.094	0.092	0.436	0.416	20.1	0.996
	(0.008)	(0.014)	(0.001)	(0.034)					18%	1%	82%	5%	(15)	0.168

Table 3 continued

	Regression coefficients								Variance components				Fit measures	
	Education		Age	Male	Father's education	Mother's education	Father manual	Material situation	Family level	Expl.[a]	Individual level	Expl.[b]	Chi2 (df)	GFI P-value
	Family	Individual												
Model B3	0.022	0.018	−0.009	0.061	0.007	−0.003	0.041	0.047	0.094	0.089	0.434	0.39	31.0	0.996
	(0.008)	(0.022)	(0.001)	(0.034)	(0.007)	(0.008)	(0.041)	(0.033)	18%	6%	82%	5%	(23)	0.124

Model A: OLS regression with robust standard errors corrected for clustering within sibling pairs; Model B: sibling models

[a] Percentage of the variance at the family level that is explained by the measured family characteristics

[b] Percentage of the variance at the individual level that is explained by the measured individual characteristics

0.066 (virtually the same as in Model A1). This model has a similar fit but uses one degree of freedom less compared to the unconstrained model, suggesting there is hardly any family bias in the association between education and self-assessed health if standard family factors are taken into account.

The qualitative results for the Dutch sample are quite similar to those of the US sample. The educational differences in health appear to be smaller in the Netherlands, but also among the Dutch siblings there is a significant and positive association between self-assessed health and education. Again, about one-fifth (18%) of the variation in self-assessed health can be attributed to the family of origin. Education, age and sex explain less of the total variation in the Dutch sample than in the US sample. This holds true for both the family level as well as the individual level. Parental education, occupation and home ownership account for a small proportion of the total family effect, namely about 6%. Comparing the coefficient for education in Model A2 with the family coefficient for education in Model B3 shows whether not taking into account the total family effect causes a bias in the effect of education. The educational coefficients in all five models are very similar and do not differ significantly from one another. Thus, there is no indication that there is a family bias in the association between education and self-assessed health in the Dutch sample.

6 Conclusions

The aim of this paper was (1) to estimate the total family effect on self-assessed health (2) to determine what proportion of the family effect is tapped by measured socio-economic variables of the family of origin, and (3) to assess to what extent the association between education and self-assessed health is biased if family factors are not taken into account.

To answer these questions, sibling data from the US and the Netherlands were used. Overall, the sibling models showed that there is more within-family variance than between-family variance in self-assessed health. In other words, individual differences seem to be much more important than differences between families. About 20% of the variation in self-assessed health is related to family factors, including all genetic, material and cultural factors (both measured and unmeasured). The remaining 80% of the variation in self-assessed health has to be attributed to purely individual differences. The total family effect on self-assessed health appears to be of equal magnitude in the two samples.

Standard indicators for the family of origin, father's occupational class (manual or non-manual) and parental education, appear to explain only a small part of the total family effect. For the US sample one-eighth of the family variance in adult self-assessed health can be explained by parental education, father's manual/non-manual class and financial well-being. Another one-eighth of the family variation is due to differences in age, sex and educational attainment of the siblings. In the Dutch sample, only 6% of the variation at the family level could be explained by measured family factors. It is important to note that in many data sets more indicators of the family's socio-economic position are available and that the indicators used in this study can be measure in much more detail. It remains an open question whether using more detailed family indicators would make a substantial difference.

No evidence was found for a substantial 'family bias' in the association between education and self-assessed health. This suggests that conventional analyses can be used to estimate largely 'unbiased effects' of education on self-assessed health if measured indicators of the family of origin are controlled for. This strengthens the conclusions from previous studies that seek to explain the association between education and self-assessed health.

At the same time, these results increase the need to come up with better explanations for the association between education and self-assessed health. And the 20% variation due to family factors calls for new studies to better understand the family of origin effects on adult health. Such studies should use twin and sibling designs as well as longitudinal individual designs. Which factors could be responsible for the unexplained part of the family effect on self-assessed health? A first logical candidate seems a genetic factor (or rather: genetic factors). Another candidate would be family culture, especially family habits related to health and health related behaviour. Although health related behaviour is correlated with the socio-economic position of the family it is likely that families develop specific habits and health beliefs that are largely independent of their socio-economic position.

This study found higher sibling correlations for education than for adult self-assessed health, suggesting family factors are more important for educational attainment. Why would shared family factors be less important for adult health? One possible explanation is that during the period when educational attainment is 'determined' for most people, brothers and sisters live at home with (one of their) parents and the environment to which they are exposed is quite similar. One's health condition, on the other hand, is shaped continuously over the life course. During adult life, siblings do not share the family home and environment anymore. The circumstances siblings life in become more divergent during adulthood. This would explain lower sibling resemblance later in life for outcomes that have important environmental, social and cultural determinants.

Some limitations of the study need to be discussed as well. The samples were quite small and prior studies (i.e., twin studies on self-assessed health) showed mixed results. Therefore, more definite conclusions can only be drawn after replication of the current findings. Replications for different age groups and twin registries are especially important. An important limitation of the US sample is its relative small size, and low participation of male siblings. Replications with larger sibling sets, should also address sex differences in more detail.

Another limitation is that the results of this study pertain to self-assessed health only. Family resemblance could differ by health indicator. I am not aware of studies showing differences in sibling (or twin) correlations for the various types of health indicators used in social epidemiological research. Still, it seems reasonable to expect that such differences exist and, hence, that one should be careful in generalizing the findings of this study.

The main strength of this study should also be stressed: the sibling design. This is a unique design that allows us to take into account the total family effect. Of course the drawback of the sibling design is that the family effect can not be separated into a genetic and non-genetic part. However, it is important to note that a sibling design is much easier to implement in large scale survey studies than a twin design. Given the substantial family impact on self-assessed health, using sibling studies and twin data to increase our knowledge on social health variations seems highly relevant. Currently, such data are only available for a limited number of countries. There is a need for more and richer data sets based on sibling and twin designs. Generating such data in more countries would not only allow us to better test the general hypotheses addressed in this study, but it would also contribute to our understanding of how social variations in health emerge in different types of societies.

References

Antonovsky, A. (1967). Social class, life expectancy and overall mortality. *Milbank Memorial Fund Quaterly,45*(2), 31–73.

Banks, J., Marmot, M., Oldfield, Z., & Smith, J. P. (2006). Disease and disadvantage in the United States and in England. *Journal of the American Medical Association,295*(17), 2037–2045.

Ben-Shlomo, Y., & Kuh, D. (2002). A life course approach to chronic disease epidemiology: conceptual models, empirical challenges and interdisciplinary perspectives. *International Journal of Epidemiology,31*(2), 285–293.

Benyamini, Y., & Idler, E. (1999). Community studies reporting association between self-rated health and mortality: Additional studies 1995–1998. *Research on Aging,21*(3), 392–401.

Blane, D., Hart, C. L., Smith, G. D., Gillis, C. R., Hole, D. J., & Hawthorne, V. M. (1996). Association of cardiovascular disease risk factors with socioeconomic position during childhood and during adulthood. *British Medical Journal, 313*(7070), 1434–1438.

Blau, P. M., & Duncan, O. D. (1967). *The American occupational structure*. New York: Wiley.

Brim, O. G., Ryff, C. D., & Kessler, R. C. (2004). The MIDUS national survey: An overview. In O. G. Brim, C. D. Ryff, & R. C. Kessler (Eds.), *How healthy are we?: A national study of well-being at midlife* (pp. 1–36). Chicago: University of Chicago Press.

Christensen, K., Holm, N. V., McGue, M., Corder, L., & Vaupel, J. W. (1999). A Danish population-based twin study on general health in the elderly. *Journal of Aging and Health,11*(1), 49–64.

Dykstra, P. A., Kalmijn, M., Knijn, T. C. M., Komter, A. E., Liefbroer, A. C., & Mulder, C. H. (2005). *Codebook of the Netherlands kinship panel study*. Den Haag: NIDI.

Erikson, R. (2001). Why do graduates live longer? Education, occupation, family and mortality during the 1990s. In J. O. Jonsson & C. Mills (Eds.), *Cradle to grave. Life-course change in modern Sweden* (pp. 211–227). Durham: Sociology Press.

Ferraro, K. F., & Farmer, M. M. (1999). Utility of health data from social surveys: Is there a gold standard for measuring morbidity? *American Sociological Review,64*(2), 303–315.

Graham, H. (2002). Building an inter-disciplinary science of health inequalities: The example of lifecourse research. *Social Science & Medicine,55*(11), 2005–2016.

Griliches, Z. (1979). Sibling models and data in economics: Beginnings of a survey. *Journal of Political Economy,87*, s37–s64.

Harris, J. R., Pedersen, N. L., McClearn, G. E., Plomin, R., & Nesselroade, J. R. (1992). Age differences in genetic and environmental influences for health from the Swedish adoption/twin study of aging. *Journal of Gerontology: Psychological Sciences,47*(3), P213–P220.

Hauser, R. M., & Mossel, P. A. (1985). Fraternal resemblance in educational attainment and occupational status. *American Journal of Sociology,91*(3), 650–673.

Hertzman, C., Power, C., Matthews, S., & Manor, O. (2001). Using an interactive framework of society and lifecourse to explain self-rated health in early adulthood. *Social Science and Medicine,53*(12), 1575–1585.

Hout, M., & Diprete, T. A. (2006). What we have learned: RC28's contributions to knowledge about social stratification. *Research in Social Stratification and Mobility,24*(1), 1–20.

Huisman, M., Kunst, A. E., Bopp, M., Borgan, J. K., Borrell, C., Costa, G., et al. (2005). Educational inequalities in cause-specific mortality in middle-aged and older men and women in eight western European populations. *Lancet,365*(9458), 493–500.

Idler, E. L., & Benyamini, Y. (1997). Self-rated health and mortality: A review of twenty-seven community studies. *Journal of Health and Social Behavior,38*(1), 21–37.

Jöreskog, K. G., & Sörbom, D. (1993). *Lisrel 8: Structural equation modelling with the SIMPLIS command language*. Chicago: Scientific Software International.

Lee, T. R., Mancini, J. A., & Maxwell, J. W. (1990). Sibling relationships in adulthood: Contact patterns and motivations. *Journal of Marriage and Family,52*, 431–440.

Leinonen, R., Kaprio, J., Jylha, M., Tolvanen, A., Koskenvuo, M., & Heikkinen, E. (2005). Genetic influences underlying self-rated health in older female twins. *Journal of the American Geriatric Society, 53*(6), 1002–1007.

Lichtenstein, P., Harris, J. R., Pedersen, N. L., & McClearn, G. E. (1993). Socioeconomic status and physical health, how are they related? An empirical study based on twins reared apart and twins reared together. *Social Science & Medicine,36*(4), 441–450.

Lundberg, O. (1993). The impact of childhood living conditions on illness and mortality in adulthood. *Social Science and Medicine,36*(8), 1047–1052.

Lynch, J. W. (2001). Socioeconomic factors in the behavioral and psychosocial epidemiology of cardio-vascular disease. In N. Schneiderman, M. A. Speers, J. A. Silva, H. Tomes, & J. H. Gentry (Eds.), *Integrating behavioral and social sciences with public health*. Washington: American Psychological Association.

Lynch, J. W., & Kaplan, G. A. (1997). Understanding how inequality in the distribution of income affects health. *Journal of Health Psychology, 2*(3), 297–314.

Marmot, M. G., & Wilkinson, R. G. (2005). *Social determinants of health*. London: Oxford University Press.

McHorney, C. A. (2000). Concepts and measurement of health status and health-related quality of life. In G. L. Albrecht, R. Fitzpatrick, & S. Scrimshaw (Eds.), *Handbook of social studies in health and medicine* (pp. 339–358). Thousand Oaks: Sage.

Mossey, J. M., & Shapiro, E. (1982). Self-rated health: A predictor of mortality among the elderly. *American Journal of Public Health,72*(8), 800–808.

Phelan, J. C., Link, B. G., Diez-Roux, A., Kawachi, I., & Levin, B. (2004). "Fundamental causes" of social inequalities in mortality: A test of the theory. *Journal of Health and Social Behavior,45*(3), 265–285.

Power, C., Matthews, S., & Manor, O. (1998). Inequalities in self-rated health: Explanations from different stages of life. *Lancet,351*(9108), 1009–1014.

Rogers, W. H. (1983). Regression standard errors in clustered samples. *Stata Technical Bulletin,13*, 19–23.

Romeis, J. C., Heath, A. C., Xian, H., Eisen, S. A., Scherrer, J. F., & Pedersen, N. L. (2005). Heritability of SF-36 among middle-age, middle-class, male-male twins. *Medical Care, 43*(11), 1147–1154.

Romeis, J. C., Scherrer, J. F., Xian, H., Eisen, S. A., Bucholz, K., & Heath, A. C. (2000). Heritability of self-reported health. *Health Service Research, 35*(5 Pt 1), 995–1010.

Roysamb, E., Tambs, K., Reichborn-Kjennerud, T., Neale, M. C., & Harris, J. R. (2003). Happiness and health: Environmental and genetic contributions to the relationship between subjective well-being, perceived health, and somatic illness. *Journal of Personality and Social Psychology,85*(6), 1136–1146.

Shavit, Y., & Blossfeld, H.-P. (Eds.). (1993). *Persitent inequality: Changing educational attainment in thirteen countries*. Boulder: Westview Press.

Sieben, I. (2001). *Sibling similarities and social stratification : the impact of family background across countries and cohorts*. Amsterdam: Thela thesis [dissertation University of Nijmegen].

Silventoinen, K., Posthuma, D., Lahelma, E., Rose, R. J., & Kaprio, J. (2007). Genetic and environmental factors affecting self-rated health from age 16–25: A longitudinal study of finnish twins. *Behavior Genetics, 37*(2), 326–333.

Stronks, K., van der Mheen, H., Looman, C. W. N., & Mackenbach, J. P. (1996). Behavioural and structural factors in the explanation of socio-economic inequalities in health: An empirical analysis. *Sociology of Health and Illness, 18*(5), 653–674.

Van de Mheen, H., Stronks, K., van den Bos, J., & Mackenbach, J. P. (1997). The contribution of childhood environment to the explanation of socio-economic inequalities in health in adult life: A retrospective study. *Social Science & Medicine,44*(1), 13–24.

Van Eijck, K. (1996). Family and opportunity: A sibling analysis of the impact of family background on education, occupation, and consumption. Tilburg: Tilburg University Press [dissertation University of Tilburg].

Van Lenthe, F. J., Gevers, E., Joung, I. M., Bosma, H., & Mackenbach, J. P. (2002). Material and behavioral factors in the explanation of educational differences in incidence of acute myocardial infarction. The globe study. *Annals of Epidemiology, 12*(8), 535–542.

Von dem Knesebeck, O., Verde, P. E., & Dragano, N. (2006). Education and health in 22 European countries. *Social Science & Medicine,63*(5), 1344–1351.

Wadsworth, M. (1997). Health inequalities in the life course perspective. *Social Science and Medicine, 44*(6), 859–869.

Wadsworth, M. E. (1999). Early life. In M. Marmot & R. G. Wilkinson (Eds.), *Social determinants of health* (pp. 44–63). Oxford: Oxford University Press.

Williams, R. L. (2000). A note on robust variance estimation for cluster-correlated data. *Biometrics,56*, 645–646.