Citation for published version:

Washbrook, E, Gregg, P & Propper, C 2014, 'A decomposition analysis of the relationship between parental income and multiple child outcomes', *Journal of the Royal Statistical Society: Series A - Statistics in Society*, vol. 177, no. 4, pp. 757-782. <https://doi.org/10.1111/rssa.12074>

DOI:

[10.1111/rssa.12074](https://doi.org/10.1111/rssa.12074)

Publication date:

2014

Document Version

Peer reviewed version

[Link to publication](https://doi.org/10.1111/rssa.12074)

This is the peer reviewed version of the following article: Washbrook, E. , Gregg, P. and Propper, C. (2014), A decomposition analysis of the relationship between parental income and multiple child outcomes. *J. R. Stat. Soc. A*, 177: 757-782, which has been published in final form at <https://doi.org/10.1111/rssa.12074>. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving.

University of Bath

Alternative formats

If you require this document in an alternative format, please contact:
openaccess@bath.ac.uk

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Appendix 1

Table A1. Variables used in analysis

Variable	Description/comments
	Child outcomes ^a
IQ	Raw total IQ score from the Wechsler Intelligence Scale for Children (WISC-III UK; Wechsler, Golombok and Rust 1992). The WISC-III UK was at the time the most up-to-date version of the WISC, the most widely used individual ability test world-wide. The short form of the scale was administered to the children at age 8 during a clinical assessment visit by ALSPAC's psychology team. The total score is derived as the sum of scores on five verbal sub-tests: information, similarities, arithmetic, vocabulary and comprehension; and five performance sub-tests: picture completion, coding, picture arrangement, block design and object assembly.
KS1 (School performance)	Derived from standardized national Key Stage 1 tests administered to all children in public schools at the end of Year 3 (when most children are aged 7). The three sub-tests cover reading, writing and mathematics. Attainment levels for the sub-tests are converted to points using the guidelines provided by the Department for Education and Skills and averaged.
Locus of control	Taken from the shortened version of the Nowicki-Strickland Internal-External scale (NSIE scales) for preschool and primary children (Nowicki and Duke 1974a). The scale consists of 12 questions read out to the child by an examiner during an ALSPAC clinical assessment visit at age 8, each requiring a yes/no answer. Responses were coded 0 or 1 and summed to create a total score. $\alpha = 0.47$.
Self esteem	Taken from the 12-item shortened form of Harter's Self Perception Profile for Children (Harter 1985). The scale was administered during the ALSPAC clinical assessment visit at age 8. Items were scored from 1 to 4 and summed to give a total. $\alpha = 0.75$.
Behaviour	Taken the Strengths and Difficulties Questionnaire (SDQ; Goodman 1997). This instrument has been shown to be a good predictor of conduct, emotional, hyperactivity and any psychiatric disorders in children of the age examined here (Goodman et al. 2000). Completed by teachers in Year 3. The SDQ comprises 4 sub-scores, each derived from the responses to 5 questions, relating to hyperactivity, emotional symptoms, conduct problems and peer problems. Items are scored from 0 to 2 and summed to create a total behaviour score. $\alpha = 0.87$.
Fat mass	Total body fat mass in grams adjusted for age of child in months, sex, height and height squared. Direct measures of the fat mass of children obtained at an ALSPAC clinical assessment visit at age 9, using dual-energy X-ray absorptiometry (DXA), a highly accurate method involving a full body scan (Morrison et al. 1994).
	Income
Log real equivalized weekly disposable household income at 33 and 47 months	Constructed from banded information on weekly disposable household income at child ages 33 and 47 months. Median values for the bands imputed using data on a comparable sample from the nationally representative Family Expenditure Survey. Converted to real values using the 1995 RPI as a base and equivalized using the OECD modified scale. We also impute the value of housing benefit for families who do not directly receive housing payments. Variable is average of two measures.

	Distal factors
Household demographics	
Family structure	Dummy equal to 1 if the mother did not live with a partner at any of 4 dates between birth and 47 months
Siblings	Dummy equal to 1 if younger sibling present in the household at 47 months. Number of older siblings in the household at 47 months (top-coded at 3).
Mother's age at birth	Mother's age at the birth of the study child
Labour market status	
Mother's employment	Mother ever worked full-time between birth and 47 months; worked part-time only; did not work at all.
Father's employment	Mother's partner was out of work never; once; more than once at any of seven dates between the pregnancy and 47 months.
Mother's and father's occupations	Defined from information gathered during pregnancy on current or last job. Responses are coded from 1 to 6 using OPCS job codes: 1 = professional; 2 = managerial/technical; 3 = skilled non-manual; 4 = skilled manual; 5 = semi-skilled; 6 = unskilled.
Education	
Mother's and father's qualifications	Collected during pregnancy. Coded: 1 = CSE/no qualifications; 2 = Vocational/O-level; 3 = A-level; 4 = Degree.
Grandparents' qualifications	Separate variables for maternal grandmother and grandfather. Coded as parental qualifications.
Neighbourhood	
Local deprivation	Rank of the Index of Multiple Deprivation (IMD) for the local electoral ward (around 5500 persons) of residence of the child at birth. The IMD is derived from 6 composite indicators in the domains of Income; Employment; Health Deprivation and Disability; Education, Skills and Training; Housing; and Geographical Access to Services. See http://www.communities.gov.uk/documents/citiesandregions/pdf/131306.pdf
Housing tenure	Always in owner-occupied accommodation between birth and 33 months; ever in public housing between birth and 33 months; other.
Ethnicity	Child is non-white.
	Proximal factors
Maternal psychosocial functioning	
Maternal anxiety/depression	Crown-Crisp Experiential Index (CCEI; Crown and Crisp, 1979). Self-assessed questionnaire completed by mothers at 6 dates spanning the early period of pregnancy to 33 months. Variable is average of the 6 measures.
Life event shocks	Derived from questions on whether each of 41 life events had occurred in recent months, and if so, how strongly the mother was affected. Responses scored on a 5-point scale from 0 (did not happen) to 5 (affected me a lot) and summed. Measure is the average score over four dates between 8 and 47 months post-birth, and so captures both the frequency and severity of shocks. Example items: A friend or relative was ill; You had problems at work; You argued with your partner; You moved house; You had a major financial problem.
Subjective financial distress	Constructed from responses to five items asking how difficult the mother currently finds it to afford food, clothing, heating, rent or mortgage and things she needs for the child. Responses scored from 0 (not difficult) to 3 (very difficult) and summed. Measure is the average of three scores taken at child age 8, 21 and 33 months.
Quality of parental relationship	Three sub-scores, all derived from information in early childhood. Affection score: derived from responses to 6 items related to how frequently

	<p>the mother and partner engage in behaviours like kissing or hugging, making plans and talking over their feelings. Responses scored from 0 to 3, summed and averaged over 3 dates between 8 and 47 months post-birth.</p> <p>Aggression score: derived from questions on how frequently the mother and her partner argued in the past 3 months, and whether 5 events such as hitting, throwing things and walking out of the house in anger occurred in the same period. Total score ranges from 0 to 14, final variable is an average over scores at 8 and 33 months.</p> <p>Shared activities score: constructed in a similar manner, using 5 items (each scored from 0 to 3) on how frequently the parents took part in activities together such as going out for a drink, a meal or to the cinema in the last 3 months.</p>
Frequency of smacking at 3	How often the mother smacks the child when he or she is naughty at 33 months. Responses are scored 1 (never), 2 (once a month or less), 3 (once a week), or 4 (daily).
Maternal social networks	<p>Two sub-scales, each administered during pregnancy and again at 21 months.</p> <p>Social networks: derived from 10 items, each scored from 0 to 3, relating to the number and strength of the mother's relationships with friends and relatives.</p> <p>Social support: same format, but here the 10 items relate to perceived levels of emotional, financial and instrumental support. Measures are average scores for each scale over the two time points.</p>
Maternal locus of control	Measured using the Adult Nowicki-Strickland Internal-External scale (ANSIE; Nowicki and Duke, 1974b). 12-item scale completed by mothers during the pregnancy.
Pre-school childcare	
Birth to age 3	<p>Six separate categorical variables indicating care by: the father; another relative or friend; nannies and babysitters; child minders; centre-based care; and other.</p> <p>Childcare mode was recorded at 8 weeks, 8, 15 and 24 months. For each type, we distinguish whether it was used at any date and if so, whether it was ever used for more than 15 hours per week.</p>
Age 3 to school entry	<p>Six separate categorical variables indicating care by: relatives (including the father); nannies; child minders; playgroups; nurseries; and other modes of care.</p> <p>Categories: not used at all; 15 hours a week or less; more than 15 hours a week.</p>
Health & health behaviours	
Health at birth	<p>Birth weight in kilograms.</p> <p>Indicators for child was born pre-term (<37 weeks gestation); whether the child was low birth weight (< 2.5 kg) but not pre-term.</p>
Smoking	<p>Mother smoked at all in pregnancy.</p> <p>Smoker in the child's household at age 4.</p>
Breast feeding	Never; < 3 months; 3 to 6 months; 6 to 12 months; > 12 months.
Eating patterns at 3	<p>Derived from a mother-completed postal questionnaire on the child's consumption of 43 different foodstuffs at 38 months. Four 'dietary types' constructed by North et al. (2000) using principal components analysis.</p> <p>Junk food: loads heavily on convenience foods such as french fries, burgers, fried foods and takeaway meals and on foods like chips, candy, cookies, chocolate and carbonated drinks.</p> <p>Healthy food: loads on vegetables, salad, fruit, fish, rice, pasta and pulses and on vegetarian substitutes for meat products.</p> <p>Traditional food: represents the traditional British 'meat and two veg' diet, loading heavily onto consumption of meat and poultry, potatoes, root vegetables, green vegetables and legumes.</p>

Snack food: relates in general to foods that require little cooking, such as puddings, cakes, cheese, bread and fruit.

Note: the purpose of these dietary types is to provide a summary of the child's eating patterns in general: they are not designed to measure specific factors such as calorie or fat content directly.

Home learning environment

Books and toys

Age at which the child is first recorded as owning at least 10 books: 6, 18, 30 or 42 months, or not at all by 42 months. (Only 6% of children own less than 10 books by age 3.)

Toy score: derived from the number of 12 different toys (such as blocks, jigsaws and interlocking toys) the child has at age 2.

Maternal teaching

Derived from questions on whether the mother teaches the child each of 10 items such as numbers, shapes, rhymes and the alphabet. Items are scored from 0 to 3 according to whether the child is first taught: not at all, by 42 months, by 30 months or by 18 months, then summed. (Lack of variation in the age 3 teaching items means that this method better distinguishes the experiences of children.)

Maternal and paternal reading/singing

Frequency the parent reads to and sings songs to the child at 18 and 42 months. Each scored from 0 to 8 (score of 8 indicates that the parent reads and sings to the child every day). Set to zero...

Trips to library, museums, etc.

Frequency child is taken to visit the library and other places of interest such as museums between the ages of 18 and 42 months. Score for each ranges from 0 to 6 (6 indicating visits of at least once a week at all three dates of measurement).

Physical home environment

Car ownership

Household without use of car at 8, 21 or 33 months.

Has garden

Household without access to garden or yard at 8, 21 or 33 months.

Noise

Noise from inside of outside household is a serious problem at 21 or 33 months.

Crowding

Average number persons per room in household at 8, 21 and 33 months.

Damp/ condensation/ mould

Damp, condensation or mould is a serious problem at 8, 21 or 33 months.

School fixed effects

School dummies

Defined only for cases in which at least 5 non-missing values of the outcome measure are observed within the same school. Children in schools with less than 5 valid observations are grouped together in a single category, and children whose school IDs are missing are similarly grouped in a separate single category. Over 70 percent of the sample children are in schools with at least 20 valid observations, and 40 to 50 percent are in schools with at least 50 valid observations.

a. Outcomes are normalized on child sex, cohort year and month of birth, and then standardized to mean 100 standard deviation 10 on the maximum available sample. The locus of control, behaviour and fat mass measures are such that higher scores indicate more adverse outcomes. Coefficients associated with these outcomes are reversed in sign in the presentation of results.

References for Table A1

- Crown S. and A.H. Crisp. 1979. *Manual of the Crown-Crisp Experiential Index*. London: Hodder & Stoughton
- Goodman, R. 1997. "The Strengths and Difficulties Questionnaire: a research note." *Journal of Child Psychology and Psychiatry* 38:581-586.
- Goodman, R., D. Renfrew, D and M. Mullick. 2000. "Predicting type of psychiatric disorder from Strengths and Difficulties Questionnaire (SDQ) scores in child mental health clinics in London and Dhaka." *European Child and Adolescent Psychiatry* 9:129-134.
- Harter, S. 1985. *Self-perception profile for children*. University of Denver.
- Morrison, J. A., P.R. Koury, B.N. Campagne, W.M.C. Chumlea, B. Specker and S.S. Guo. 1994. "Body-Composition Measures from Underwater Weighing and Dual-Energy X-Ray Absorptiometry in Black-and-White Girls - a Comparative-Study." *American Journal of Human Biology* 6:481-490.
- North, K., P. Emmett and A.S. Team. 2000. "Multivariate Analysis of Diet among Three-Year-Old Children and Associations with Socio-Demographic Characteristics." *European Journal of Clinical Nutrition* 54:73-80.
- Nowicki, S. and M.P. Duke. 1974a. "A preschool and primary internal-external control scale." *Developmental Psychology* 10:874-881.

- Nowicki, S. and M.P. Duke.. 1974b. "A locus of control scale for college as well as non-college adults." *Journal of Personality Assessment* 38: 136-137.
- Wechsler, D., S. Golombok and J. Rust. 1992. *WISC-IIIUK Wechsler Intelligence Scale for Children – Third Edition UK Manual*. Sidcup, UK: The psychological Corporation.

Table A2. Descriptive statistics for all variables

Variable	Mean	Std. Dev.	Min	Max	Fraction non-missing ^a
Child outcomes ^b					
IQ	100.18	9.92	66.02	129.00	0.60
Key Stage 1 (KS1)	100.82	9.27	60.86	124.37	0.92
Locus of control	99.89	9.98	70.36	125.25	0.57
Self esteem	100.03	9.98	59.37	119.79	0.62
Behavior at 7	99.90	9.93	85.27	152.10	0.35
Fat mass	99.91	9.83	74.83	158.35	0.65
Income					
Average weekly disposable income ^c	222.54	100.32	34.99	625.57	1.00
Distal variables					
Household demographics					
Family structure: single mother by 47 months	0.14	0.35	0	1	1.00
Siblings: younger sibling by 47 months	0.41	0.49	0	1	0.92
Siblings: number of older siblings at 47 months	0.95	1.11	0	3	0.92
Mother's age at birth	28.57	4.70	15	44	1.00
Child is non-white	0.04	0.19	0	1	0.95
Labour market status					
Mother's employment: not employed pre-school	0.25	0.43	0	1	0.87
Mother's employment: part-time only pre-school	0.55	0.50	0	1	0.87
Mother's employment: full-time pre-school	0.20	0.40	0	1	0.87
Father's employment: always in work to age 4	0.81	0.39	0	1	0.66
Father's employment: out of work at 1 date by age 4	0.10	0.30	0	1	0.66
Father's employment: out of work > 1 date by age 4	0.09	0.29	0	1	0.66
Mother's occupation	2.85	1.07	1	6	0.81
Partner's occupation	3.00	1.31	1	6	0.88
Education					
Mother's qualifications	2.37	0.92	1	4	0.97
Partner's qualifications	2.44	1.04	1	4	0.94
Grandmother's qualifications	1.76	0.78	1	3	0.72
Grandfather's qualifications	2.00	1.02	1	4	0.68
Neighbourhood					
Local deprivation: Rank of ward IMD score	4536	2523	0	8379	0.91
Housing tenure: owner-occupied birth to 33 months	0.75	0.43	0	1	0.99
Housing tenure: ever in social housing	0.16	0.37	0	1	0.99
Housing tenure: other birth to 33 months	0.09	0.28	0	1	0.99
Proximal variables					
Maternal psychosocial functioning					
Maternal depression/anxiety	11.51	6.26	0	41.50	1.00
Life event shocks	2.87	1.91	0	18.92	1.00
Subjective financial distress	3.12	3.20	0	15	0.99
Quality of parental relationship: affection	14.99	2.88	0	18	0.96
Quality of parental relationship: aggression	3.92	2.82	0	14	0.93
Quality of parental relationship: shared activities	5.98	2.54	0	15	0.93
Frequency of smacking at age 3	2.29	0.72	1	4	0.87
Maternal social networks sub-score	23.39	3.64	3	29	0.98
Maternal social support sub-score	20.28	4.62	1.5	30	0.96
Maternal locus of control	4.20	2.13	0	12	0.83

Variable	Mean	Std. Dev.	Min	Max	Fraction non-missing ^a
Pre-school childcare					
Birth to age 3 (dummy variables)					
Partner 1-15 hrs pwk	0.49	0.50	0	1	0.99
Partner > 15 hrs pwk	0.42	0.49	0	1	0.99
Friend/relative 1-15 hrs pwk	0.33	0.47	0	1	1.00
Friend/relative > 15 hrs pwk	0.08	0.27	0	1	1.00
Child minder 1-15 hrs pwk	0.06	0.24	0	1	1.00
Child minder > 15 hrs pwk	0.11	0.31	0	1	1.00
Nanny 1-15 hrs pwk	0.06	0.23	0	1	1.00
Nanny > 15 hrs pwk	0.05	0.21	0	1	1.00
Nursery 1-15 hrs pwk	0.08	0.28	0	1	1.00
Nursery > 15 hrs pwk	0.04	0.19	0	1	1.00
Other 1-15 hrs pwk	0.02	0.15	0	1	1.00
Other > 15 hrs pwk	0.01	0.09	0	1	1.00
Age 3 to school entry					
Partner/friend/relative 1-15 hrs pwk	0.15	0.35	0	1	0.89
Partner/friend/relative > 15 hrs pwk	0.06	0.23	0	1	0.89
Child minder 1-15 hrs pwk	0.06	0.23	0	1	0.89
Child minder > 15 hrs pwk	0.05	0.22	0	1	0.89
Nanny 1-15 hrs pwk	0.02	0.13	0	1	0.89
Nanny > 15 hrs pwk	0.02	0.15	0	1	0.89
Playgroup 1-15 hrs pwk	0.35	0.48	0	1	0.89
Playgroup > 15 hrs pwk	0.01	0.12	0	1	0.89
Nursery 1-15 hrs pwk	0.32	0.47	0	1	0.89
Nursery > 15 hrs pwk	0.13	0.34	0	1	0.89
Other 1-15 hrs pwk	0.05	0.22	0	1	0.89
Other > 15 hrs pwk	0.19	0.39	0	1	0.89
Health & health behaviors					
Birth weight (kg)	3.42	0.55	0.65	5.64	0.99
Gestation < 37 weeks	0.05	0.22	0	1	1.00
Low birth weight (<2.5kg) and not pre-term	0.02	0.14	0	1	0.99
Mother smoked in pregnancy	0.25	0.43	0	1	0.95
Smoker in household age 4	0.38	0.49	0	1	0.89
Never initiated breastfeeding	0.25	0.43	0	1	0.98
Breastfed < 3 months	0.24	0.43	0	1	0.98
Breastfed 3-6 months	0.17	0.37	0	1	0.98
Breastfed 6-12 months	0.24	0.43	0	1	0.98
Breastfed > 12 months	0.10	0.30	0	1	0.98
Junk food score at age 3	-0.09	2.95	-7.43	29.96	0.76
Healthy food score at age 3	-0.05	2.72	-6.46	22.80	0.76
Traditional food score at age 3	0.01	2.51	-7.93	17.01	0.76
Snack food score at age3	0.07	1.95	-9.31	12.16	0.76
Home learning environment					
Child first owned 10+ books at 6 months	0.19	0.39	0	1	1.00
Child first owned 10+ books at 18 months	0.47	0.50	0	1	1.00
Child first owned 10+ books at 30months	0.20	0.40	0	1	1.00
Child first owned 10+ books at 42 months	0.07	0.26	0	1	1.00
Child never owned 10+ books by 42 months	0.06	0.24	0	1	1.00
Toy score at age 2	15.21	2.22	1	20	0.90
Maternal teaching score	27.32	2.47	11	30	0.92
Maternal reading and singing score at 18 mths	7.02	1.42	0	8	0.94
Maternal reading and singing score at 42 mths	6.28	1.89	0	8	0.92

Variable	Mean	Std. Dev.	Min	Max	Fraction non-missing ^a
Paternal reading and singing score at 18 mths	4.55	2.28	0	8	0.90
Paternal reading and singing score at 42 mths	4.35	2.05	0	8	0.86
Outings to library score	1.42	1.60	0	6	0.83
Outings to museums/places of interest score	2.01	1.86	0	6	0.84
Physical home environment					
Without access to car at 8, 21 or 33 months	0.11	0.32	0	1	0.99
Without access to garden at 8, 21 or 33 months	0.07	0.26	0	1	0.99
Noise serious problem at 21 or 33 months	0.08	0.26	0	1	0.98
Average crowding index (persons per room)	0.79	0.31	0.19	9.0	0.98
Damp/mould/condensation ever serious problem	0.13	0.33	0	1	0.99
Average within-school scores^d					
Average IQ in school	99.42	2.61	88.55	107.62	0.48
Average Key Stage 1 score in school	100.48	2.57	65.23	108.35	0.79
Average locus of control score in school	100.36	2.16	89.57	108.17	0.46
Average self esteem score in school	99.97	1.73	92.19	106.57	0.50
Average behavior score in school	99.86	2.38	90.65	110.79	0.26
Average fat mass score in school	100.10	1.62	94.28	108.11	0.52

Notes. Variables with minimum of 0 and maximum of 1 are dummy variables. See Appendix Table A1 for variable definitions.

- Statistics defined over the full sample with non-missing income and at least one child outcome measure (N = 9476).
- Outcomes are standardized to mean 100, standard deviation 10 on the full sample of observations available. Differences in the mean and standard deviation of the working samples are due to the dropping of cases with missing household income. For locus of control, behavior and fat mass higher scores indicate more adverse outcomes.
- Equalized. 1995 prices. Measure is logged in regression analysis.
- Defined only for children with at least 4 other non-missing peers' scores. For illustrative purposes only, school dummies are used in multivariate analysis.

Table A3. Correlations between outcome measures

	IQ	KS1	Locus of control	Self esteem	Behaviour	Fat mass
IQ	1 N = 5708					
KS1	0.64 N = 5162	1 N = 8727				
Locus of control	0.33 N = 4907	0.28 N = 4883	1 N = 5390			
Self esteem	0.18 N = 5346	0.21 N = 5295	0.22 N = 5055	1 N = 5857		
Behaviour	0.29 N = 2515	0.37 N = 2959	0.14 N = 2365	0.22 N = 2567	1 N = 3294	
Fat mass	0.08 N = 4946	0.10 N = 5536	0.06 N = 4681	0.02 N = 5072	0.08 N = 2605	1 N = 6113
(p = 0.088)						

Note. All correlations are significant at the 1% level unless otherwise marked.

Table A4a. Decomposition of the contribution of household demographics

	Child outcome measure					
	Cognitive		Non-cognitive			Health
	IQ	KS1	Locus of control	Self esteem	Behaviour	Fat mass
Total contribution of household demographics : $(\gamma\beta_k + \theta_k)\alpha_k$	0.55** [100%]	0.48** [100%]	0.38* [100%]	0.15 [100%]	0.32* [100%]	-0.18 [-100%]
Of which explained by: $(\gamma_j\beta_{jk}\alpha_k)$						
Maternal psychosocial functioning	-	-	-	0.08* [55.5%]	0.12** [38.6%]	-
Pre-school childcare	-	-	-	-	-	-
Health & health behaviours	-	0.06* [11.9%]	0.06 [†] [15.0%]	-	0.09* [29.7%]	0.13** [69.0%]
Home learning environment	0.17** [30.2%]	0.31** [65.2%]	0.17** [43.7%]	0.14* [95.9%]	-	-
Physical home environment	-	-	-	-0.07 [†] [-45.2%]	-	-0.08* [-44.6%]
School fixed effects	-	-	-	-	-	-
Unexplained: $\theta_k\alpha_k$	0.25 [45.7%]	-	-	-	-	-0.24 [-130.4%]

Table A4b. Decomposition of the contribution of labour market status

	Child outcome measure					
	Cognitive		Non-cognitive			Health
	IQ	KS1	Locus of control	Self esteem	Behaviour	Fat mass
Total contribution of labour market status : $(\gamma\beta_k + \theta_k)\alpha_k$	0.79** [100%]	0.94** [100%]	0.75* [100%]	0.23 [100%]	0.66* [100%]	0.12 [100%]
Of which explained by: $(\gamma_j\beta_{jk}\alpha_k)$						
Maternal psychosocial functioning	0.09** [11.8%]	0.09** [10.0%]	0.08 [†] [10.9%]	0.09* [39.9%]	0.09 [†] [13.6%]	0.06 [†] [55.7%]
Pre-school childcare	0.07 [†] [8.2%]	-	0.12* [16.2%]	-	-	-
Health & health behaviours	-	0.09* [9.2%]	0.09* [12.4%]	-	0.18** [26.4%]	0.16** [140.9%]
Home learning environment	0.07* [8.6%]	0.06 [†] [6.6%]	-	0.05 [†] [21.0%]	-	-
Physical home environment	-	-	-	-	-	-0.07* [-64.3%]
School fixed effects	-	-	-	-	-	-
Unexplained: $\theta_k\alpha_k$	0.48* [60.8%]	0.67** [71.8%]	0.40 [53.2%]	0.20 [86.3%]	0.36 [54.2%]	-

Table A4c. Decomposition of the contribution of education

	Child outcome measure					
	Cognitive		Non-cognitive			Health
	IQ	KS1	Locus of control	Self esteem	Behaviour	Fat mass
Total contribution of education : $(\gamma\beta_k + \theta_k)\alpha_k$	3.09** [100%]	2.05** [100%]	1.60** [100%]	0.58** [100%]	0.66** [100%]	0.88** [100%]
Of which explained by: $(\gamma_j\beta_{jk}\alpha_k)$						
Maternal psychosocial functioning	0.26** [8.6%]	0.22** [10.9%]	0.21** [13.1%]	0.14** [24.2%]	0.14* [20.7%]	-
Pre-school childcare	-	-	0.14** [8.4%]	-0.08 [†] [-14.5%]	-0.13* [-19.4%]	-
Health & health behaviours	0.33** [10.8%]	0.20** [9.5%]	0.23** [14.5%]	-	0.19* [29.0%]	0.42** [48.2%]
Home learning environment	0.22** [7.1%]	0.15** [7.3%]	0.10 [6.2%]	0.10 [†] [18.0%]	-	-
Physical home environment	-	-	-	-	-	0.03* [3.2%]
School fixed effects	-	-	-	-	-	-
Unexplained: $\theta_k\alpha_k$	2.15** [69.7%]	1.51** [73.7%]	0.87** [54.2%]	0.50** [87.0%]	0.44* [67.7%]	0.35* [39.6%]

Table A4d. Decomposition of the contribution of neighbourhood

	Child outcome measure					
	Cognitive		Non-cognitive			Health
	IQ	KS1	Locus of control	Self esteem	Behaviour	Fat mass
Total contribution of neighbourhood : $(\gamma\beta_k + \theta_k)\alpha_k$	0.72** [100%]	1.15** [100%]	0.16 [100%]	0.12 [100%]	0.71** [100%]	0.55* [100%]
Of which explained by: $(\gamma_j\beta_{jk}\alpha_k)$						
Maternal psychosocial functioning	0.05 [†] [6.6%]	0.08** [6.5%]	0.08* [49.4%]	0.08** [66.1%]	0.20** [28.0%]	0.08** [15.4%]
Pre-school childcare	-	-	-	-	0.12** [17.3%]	-
Health & health behaviours	-	0.15** [13.3%]	0.14* [86.1%]	-	0.29** [40.3%]	0.19** [34.3%]
Home learning environment	0.09** [12.2%]	0.10** [9.0%]	-	0.05 [†] [44.6%]	0.12** [17.3%]	-
Physical home environment	-	-	-	-0.19 [†] [-159.5%]	-	-0.25* [-45.5%]
School fixed effects	-	-	-	-	-	-
Unexplained: $\theta_k\alpha_k$	0.51** [71.7%]	0.77** [66.7%]	-	0.11 [86.8%]	-	0.49* [90.1%]

Notes: N = 9476. Numbers in the higher row of each pair are selected path coefficients shown in equation 5. Numbers in square brackets express the coefficient as a proportion of the total contribution of the distal grouping given in the first row of each sub-table. Standard errors (not shown) calculated by non-parametric bootstrap.

** p<.01; * p<.05; [†] p<.10.

Appendix 2: Illustrative example of path model

Consider a decomposition model between income and an outcome with a single distal factor other than income (e.g. parental education) and a single proximal factor (e.g. the frequency a parent reads to the child). The raw income gradient in the outcome is 6.3 points, implying a log-point change in income is associated with a 0.63 standard deviation improvement in the outcome. The parameters of the structural equations (numbered as in Section 2.2) are presented below.

$$Y_i = 70 + 6.3(\delta)\text{Income}_i + e_i \quad (5)$$

$$Y_i = 75 + 5(\gamma)\text{Reading}_i + 3(\theta)\text{Education}_i + 2(\pi)\text{Income}_i + \mu_i \quad (1)$$

$$\text{Reading}_i = -0.15 + 0.3(\beta)\text{Education}_i + 0.5(\lambda)\text{Income}_i + \eta_i \quad (2)$$

$$\text{Education}_i = -1.5 + 0.4(\alpha)\text{Income}_i + v_i \quad (3)$$

The table below shows the way the underlying parameters can be combined to focus on different aspects of the income gradient.

Component	Formula	Calculation	Path coefficient	% of raw gradient
Total gradient	$\gamma\beta\alpha + \gamma\lambda + \theta\alpha + \pi$ $= \delta$	$(5 \times 0.3 \times 0.4) + (5 \times 0.5) + (3 \times 0.4) + 2$	6.3	100%
Decomposition 1				
A. Total proximal contribution (sum of i and ii)	$\gamma(\beta\alpha + \lambda)$	$5 \times [(0.3 \times 0.4) + 0.5]$	3.1	49%
i. Income-proximal contribution	$\gamma\lambda$	(5×0.5)	2.5	40%
ii. Distal-proximal contribution	$\gamma\beta\alpha$	$(5 \times 0.3 \times 0.4)$	0.6	9%
B. Total unexplained contribution (sum of iii and iv)	$\theta\alpha + \pi$	$(3 \times 0.4) + 2$	3.2	51%
iii. Income-unexplained contribution	π	2	2	32%
iv. Distal-unexplained contribution	$\theta\alpha$	(3×0.4)	1.2	19%
Decomposition 2				
C. Direct income contribution (sum of i and iii)	$\gamma\lambda + \pi$	$(5 \times 0.5) + 2$	4.5	72%
i. Income-proximal contribution	$\gamma\lambda$	(5×0.5)	2.5	40%
iii. Income-unexplained contribution	π	2	2	32%
D. Total distal contribution (sum of ii and iv)	$(\gamma\beta + \theta)\alpha$	$[(5 \times 0.3) + 3] \times 0.4$	1.8	28%
ii. Distal-proximal contribution	$\gamma\beta\alpha$	$(5 \times 0.3 \times 0.4)$	0.6	9%
iv. Distal-unexplained contribution	$\theta\alpha$	(3×0.4)	1.2	19%

The first decomposition focuses on the portion of the raw income gradient that can be explained by proximal factors. The total proximal contribution in row A is a measure of how far the income gradient is predicted to fall if all unconditional associations between proximal factors and income were eliminated – in this example it equates to the elimination of the correlation between income and parental reading. This decomposition focuses on the extent to which differences in the measured immediate environments of children can explain the

income gradient, abstracting from whether it is income or other distal factors that are the source of these environmental differences. Here, the greater reading in higher-income families generates a gradient of 3.1 points, or 49% of the raw gradient.

The total proximal contribution reflects two processes. First, income has a direct relationship with reading, holding education constant (λ). The path coefficient of 2.5 implies that equalising income across families while leaving education unchanged is associated with a change in reading behaviour that reduces the overall predicted gradient by 40%. This is less than the total contribution of reading behaviour because lower income families also tend to have less education (captured by α), which itself exerts an independent influence on reading behaviour (β). 0.6 points, or 9% of the overall gradient, is generated by the association of education with reading, combined with the concentration of low-educated parents in poorer families.

3.2 points, or 51%, of the raw gradient is attributed to other differences between low and higher income families besides reading behaviour. Some of these differences are conditionally associated with education – 19% of the total gradient is explained by something low- and high-educated parents do differently even when income is held constant. Finally, 32% of the raw gradient is accounted for by unobserved processes that are directly linked to income. The π coefficient implies that equalising income, while holding reading and education constant, is predicted to reduce the observed gradient by 2 points.

The second decomposition focuses on the extent to which income itself generates the observed gradient, separate from the confounding influence of other distal characteristics. The component C gives this “direct” effect of income, abstracting from whether it operates through measured proximal factors or unobserved processes. This is the type of income effect that is the focus of experimental and other causal analyses. In this example, equalising income across families, holding education constant, is predicted to lower the observed gradient by 72%. The division of this direct effect into 40% via the effect on reading behaviour and 32% via other unmeasured processes corresponds to the object of inquiry in the meditational studies of the conditional effect of income (e.g Guo and Harris 2000).

Component D, the part of the observed gradient that is due to the confounding of income with other distal factors, is the part that is normally treated as a nuisance in studies aiming to isolate the effect of income on outcomes. Inspection of this component provides evidence on the question: if it is not money itself that generates the poorer outcomes of low-income children, what is it? In this example education is the only potential confounder (accounting for the remaining 28% of the gradient), but in a realistic application education will compete with other distal characteristics such as family structure and labour market experiences. Further, each confounding distal contribution can be further decomposed into explained and unexplained components. The numbers for this example imply that a third of the overall gradient generated by education differences is due to the association between education and reading (0.6/1.8), while two-thirds is due to the association between education and unmeasured environmental influences that matter for the outcome.