UvA ECONOMETRICS

Discussion Paper: 2011/06

Childhood Intelligence and Adult Mortality in the Brabant Data Set: First Report

J.S. Cramer

www.feb.uva.nl/ke/UvA-Econometrics

Amsterdam School of Economics
Department of Quantitative Economics
Valckenierstraat 65-67
1018 XE AMSTERDAM
The Netherlands

UvA UNIVERSITEIT VAN AMSTERDAM

Childhood Intelligence and Adult Mortality in the Brabant Data Set: First Report

J. S. Cramer*
University of Amsterdam and Tinbergen Institute†

July 18, 2011

Abstract

The Brabant Data Set, now freely accessible, contains information on a sample cohort of 3 000 individuals born around 1940 from surveys in 1952, 1983 and 1993, as well as on deaths between 1994 and 2009. In line with numerous epidemiological studies we find that among the early variables recorded at age 12 the only significant determinant of adult mortality is intelligence. Preliminary attempts to trace this effect in the later surveys are not successful.

1 The Brabant data set

The Brabant data set combines information from three surveys and one archive. Data collection started in 1952 with an educational survey among a sample of schoolchildren in the sixth form of primary schools in the Dutch province of Brabant. In 1983 a postal survey dealing with education and employment history was sent to some 80% of the initial sample who could be traced, and this was repeated in 1993 with additional questions about

*I am indebted to Hans Bosma, Govert Bijwaard and Hans van Kippersluis for comments on an earlier draft.
†J.S.Cramer@uva.nl.

entrepreneurship. The 2998 individuals of the present data set have all participated in the 1952 survey and in at least one of the two subsequent surveys, with item nonresponse of varying degree. On the basis of their full names and dates of birth it was possible to find out whether they have died (and on what date) between 1994 and 2009, and this information has been added. Fuller details of the data (and their weaknesses) as well as of the method of analysis are given in a companion technical report [1]. The data set with full documentation (in part in Dutch) is available on the website www.DANS/KNAW/nl - look for *Brabantse zesdeklassers 1952-2010.*

By the initial survey, the sample represents a cohort of Brabant children born between July 1937 and June 1941, with early 1940 the median date of birth. The observation of deaths is confined to a window from October 1, 1994 to February 3, 2009, when ages in the cohort range from 53 to 72, the onset of old age. Mortality is still fairly low, and the sample records deaths of only 14% of the men and of 9% of the women; incomplete lives, which are rather uninformative about mortality, dominate the sample likelihood. The sample is therefore a tenuous basis for a study of lifetimes, and only strong relationships will show up. Five or ten years from now there will be substantially more deaths, and the documentation at DANS shows how these can be retrieved[1].

In addition to the preponderance of incomplete lives, it is a source of concern that sample deaths fall short of mortality in the correponding population cohort by 22% for men and by 16% for women. Various reasons for this shortfall have been put forward, but none is conclusive. Whether it affects (and invalidates) the present analysis is unclear.

The data set also has some merits. Since it represents a cohort we need not bother about historical changes in general conditions - all respondents have lived through the same epoch. Another advantage is that it contains direct information from the participants at age 12 as well as at ages 43 and 53.

In this first report we scan individual characteristics from the 1952 survey for their effect on mortality fifty years later. We find a strong effect of intelligence on longevity, a relationship already well established by numerous epidemiological studies. Information from the later surveys may shed some light on the underlying mechanism, but this work is still under way.

[1] Since this part of the archive involves the identification of individuals access is subject to permission, granted to *bona fide* researchers only.

2 Model and method

The effect of various determinants on mortality is established by Maximum Likelihood estimation of the proportional hazard model

$$h(t,i) = h^*(t).\exp(x_i'\beta) \qquad (1)$$

with h the hazard or age-specific mortality rate, t age, i the individual, and h^* the time dependent hazard; the second term represents the effect of fixed individual characteristics x_i. In view of the weakness of the sample information the specification of h^* and its time gradient have been borrowed from aggregate statistics. In the age range under review the baseline hazard $h^*(t)$ has the same time gradient for men and women, but different levels: women's hazards are only .57 of the men's. Men and women are therefore treated separately in all analyses. Since missing observations are omitted the number of observations varies from case to case. The maximum is the sample size, 1790 for men and 1208 for women[2].

The object of the estimations is β of (1). At times x is a logarithm, and then β is an elasticity. We gauge the significance of β by its t value; significance is established if t exceeds 2, and values exceeding 1.5 are treated as worthy of note. In the tables, estimates with a t exceeding 1.5 are set in boldface, and if exceeding 2 in boldface and given an asterisk. We report standardized hazard ratios HRs, the effect on mortality of an increase in x by one standard deviation,

$$HRs = \exp(\beta_j.sd(x_j)), \qquad (2)$$

and its 95% confidence interval, and refer to elasticities where this is appropriate. The original Maximum Likelihood estimates are given in an Appendix.

[2]This preponderance of males is due to special efforts in the 1983 survey to boost their response.

3 Early determinants of adult mortality

3.1 Family background

From the 1952 survey we know a great deal about the school results of the children and a little about their family background. The scene is set in a rural province with a predominantly Roman Catholic population, a long time ago; families with 10 or 12 children are no exception, and 95% of all children have the usual complement of two parents. The issue of parental divorce does not arise.

We have an indicator of the parental *social class*, initially a threefold classification by the father's occupation, here extended by adding a lowest class for asocial households - asocial, that is, by 1952 standards, and presumably by the judgment of the schoolmaster. The variable *working child* is a (0,1) dummy variable for children made to help in their parents' trade, often an indication of farmer families. *Education of the parents* is measured on a six (effectively five) point scale of the highest level attained.

Table 1. Effect of family background variables

variable	scale	n	HRs, ci	n	HRs, ci
		men		women	
social class	0,...3	1622	.93 (.81-1.05)	1095	1.09 (.87-1.32)
working child	0,1	1575	.98 (.84-1.11)	1058	1.08 (.88-1.28)
# of children	1,2,..	1642	**.88 (.76-1.00)**	1170	.99 (.79-1.18)
birthrank	1,2,..	1635	.94 (.81-1.06)	1172	1.08 (.87-1.28)
education father	2,..6	1268	1.07 (.92-1.21)	680	.78 (.50-1.06)
education mother	2,..6	1263	.96 (.80-1.11)	682	.93 (.65-1.22)

By the standards we have adopted none of the family background variables is significant. The one exception that deserves mention is a negative effect on mortality of the number of siblings among men. But with no effect of birth rank I do not know what to make of this.

3.2 School performance and intelligence tests

In addition to school marks for several subjects we have four general indicators of scholarly ability, viz. test scores for three kinds of intelligence test and the teacher's assesment of the child's educational prospects. We also know whether the child has earlier repeated one or more forms because of insufficient progress.

The three IQ tests deserve special consideration. They were especially designed for the 1952 survey with a view to distinguish innate abilities from acquired knowledge ([2], p.35-36); unfortunately, the actual tests have not been reported nor preserved. The first or *standard IQ is* an adaptation of an existing standard intelligence test for schoolchildren devised by Snijders [3]: it is composed in roughly equal parts of questions relating to school subjects (grammar, spelling, arithmetic, geography, history) and questions that call for reasoning and deduction (sequences, geometrical puzzles). The second, presented in the 1952 report as a test of *abstract reasoning*, is the progressive matrices test of Raven [4]. It is known that this specifically reflects the powers of abstraction and goal management, i.e. the strategy of problem solving, as well as the ability to cope with novel challenges - see Carpenter et al [5]. The third is a test of *vocabulary* In view of their skew distributions we take logarithms of the test scores, which we denote as *IQ, IQa* and *VOC*. The *teacher's assesment* is rated on a six-point scale of suitable further education, and then there is the number of *forms repeated* because of insufficicent progress. These five indicators of ability and school performance are somewhat correlated, but not to such an extent that any single one can be regarded as redundant.

Table 2 Correlations of school performance and intelligence tests

variable	scale	IQ	IQa	vocab	teacher	forms rep'd
			men (n = 1411)			
IQ	logs	1				
IQa	logs	.59	1			
vocabulary	logs	.63	.38	1		
teacher	1,2,..6	.50	.33	.48	1	
forms rep'd	0,1,2,	-.21	-.10	-.15	-.36	1
			women (n = 1055)			
IQ	logs	1				
IQa	logs	.57	1			
vocabulary	logs	.65	.37	1		
teacher	1,2,..6	.55	.35	.51	1	
forms rep'd	0,1,2,	-.26	-.19	-.19	-.29	1

The result of entering each variable in turn in the proportional hazards model is shown in Table 3.

Table 3. Effect of school performance and intelligence tests

variable	n	HRs, ci men	n	HRs, ci women
logIQ	1600	**.83*(.72-.94)**	1132	**.82 (.66-.99)**
logIQa	1463	**.82*(.70-.93)**	1125	**.78*(.62-.93)**
logVOC	1456	.96 (.83-1.09)	1121	.86 (.68-1.04)
teacher	1668	.94 (.81-1.06)	1161	.92 (.65-1.19)
forms rep'd	1690	1.00 (.88-1.13)	1178	1.07(.88-1.27)

The two school performance indicators have no effect. High test scores for all three intelligence tests reduce mortality, though in quite different degree. Abstract IQ or IQa is significant for both sexes, standard IQ is significant for men and just significant for women, and vocabulary has a much weaker and insignificant effect. In the relationship with mortality IQa is dominant: it is stronger than the others (larger t value, larger β_1 and higher loglikelihood compared to most other variables - see Appendix), and this effect is more pronounced for women than for men.

The HRs are much the same for men (.82) and women (.78): one standard error higher IQa score reduces hazards by 18% among men and 22% among women. Since we have taken logarithms of the test scores, β_1 is an elasticity. The estimates of -1.4 for men and -1.8 for women therefore indicate that a 10% higher IQa score reduces mortality by 14% among men and by 18% among women. In the age range under consideration, hazards for the sample cohort increase by 8% per year; their reduction corresponding to a 10% higher intellligence score is therefore equivalent to rejuvenation by 1.75 years for men and 2.25 years for women. This need not mean that a 10% higher intelligence at twelve confers a gain of two years lifetime, for there may well be compensating changes in mortality at older age.

Thus IQa is a the stronger determinant of mortality, stronger than IQ and far stronger than VOC. Standard IQ has a correlation of .6 with IQa, so standard IQ has hardly any additional effect on mortality beyond IQa; with a 10% higher value for both tests, the HRs is .77 for both sexes. Moreover, I have allowed for the effect of age (as apart from schooling) by adding *number of forms repeated* to IQa, but with no discernible effect. In the sequel I shall therefore consider IQa exclusively as the principal indicator of reduced adult mortality.

To complete the analysis we have examined the effect of introducing each of the family background variables of Table 1 in turn along with IQa, thus controlling for their effect. None of them obtains a significant coefficient; some (but not all) do affect the elasticity of IQa. The pattern is different between sexes. Among men, the addition of the variables family size and birth rank raises the elasticity (in absolute terms) from -1.4 to -1.7; among women, social class raises the elasticity from -1.8 to -2.6, and education of the father and of the mother reduce it to -1.2 and -1.5 respectively. It is difficult to recognize a pattern in these changes, but they do indicate that the relation of IQa to mortality is not entirely independent of the child's home environment. The relevant estimates are given in the Appendix.

3.3 Discussion

The present analysis confirms the association of higher early intelligence with reduced adult mortality that has been firmly established over the last decades by many independent studies. In the first systematic review of the field, Batty, Deary and Gottfredson [6] list nine independent studies that all arrive at the same conclusion. This has recently been superseded by the survey of Calvin et al [7], who identify 27 studies and give a meta-analysis of 16 studies that are based on independent data sets. Twelve out of the 27 refer to intelligence tests at ages between 10 and 12, when intelligence is not yet affected by differences in schooling[3], and seven of these are included in the meta-analysis. Their HRs range from .85 to .64. By this standard the present values of .82 for men and .78 for women are rather modest.

While these seven studies [8] to [14] are comparable to the present analysis, there are some differences, too. In the first place, outside Brabant intelligence is never measured by Raven's Progressive Matrices Test, but by tests like Brabant 's standard IQ test, here a weaker indication of adult mortality than the Raven PMT. Secondly, while in Brabant the effect of intelligence on mortality is stronger for women than for men, this is not always so in the other studies. Of the seven analyses, [10] has an all male sample, and [9] and [12] do take gender into account but estimate a single common HRs for men and women. Among the remaining four, [8] reports a stronger effect for women, in [13] the HRs are the same, and in [11] and [14] IQ has no significant effect on mortality for women - in [14], a higher IQ is even associated with a *higher* mortality of older women (beyond 60). All that is clear is that gender makes a difference, which must be of further interest (like the overall difference in hazards between men and women).

Deary [15] has forcefully argued that the interpretation of the principal inverse association of intelligence with mortality is a major unsolved problem. As a species man has achieved supremacy by his superior cognitive ability, and this may also hold at the individual level; but something more precise is called for. It is easy to put forward a number of intermediary aspects of adult life by which intelligence may lead to a longer life, for there is an abundance of studies showing that higher levels of education, higher standards of living, higher social class and so on all contribute to longer lifetimes, and such variables may well be associated with intelligence. In practice, however,

[3] This is the reason for preferring measurement at 12 rather than 18 or 20. It is known that early intelligence differentials persist until old age; see [16].

intelligence, education, social class and income are inextricably interrelated, and it is very difficult to disentangle them at the individual level.

4 Some intermediary variables

In a first and admittedly crude attempt we explore the role of a few key variables from the later surveys. These are *education, gross earnings* as an indication of earning power, and *net family income* and *net wealth* as measures of living standards. *Education* is the highest level attained on a scale from 3 to 6, the maximum recorded in the two surveys of 1983 and 1993 (so as to minimize nonresponse). The *earnings* and *income* variables have been constructed from a list of detailed survey questions about wages, holiday money, assistance payments and so on; but the partner's contribution to family income is obtained from a single question. In the generation under review relatively few women went out to work, so for women earnings are often a missing variable and family income a doubtful estimate. We have taken logarithms of all four money variables. *Net wealth* is measured on a scale from 1 to 10 that starts with heavy debts and ends with an open class of affluence. All these financial variables suffer from a poor response. In addition to the standard hazard ratios we also record the correlation of the determinant under review with logIQa.

The result of Table 4 is disappointing. One would have hoped for strong correlations with logIQa and strong associations with mortality for all or some of the variables, but this is not what we find. Among men, we find a single significant variable *family income '93*, and a mild association of *education* and *logearnings83* with mortality; among women, we find nothing of note. Further exploration of the intermediary surveys and the use of more sophisticated models is in order before we can conclude that the surveys of 1983 and 1993 do not report the essential variables that mediate between early intelligence and adult mortality.

Table 4. Effects of education, earnings, income and wealth.

variable men	n	HRs, ci	R(IQa)	n	HRs, ci women	R(IQa)
education	1589	**.88 (.75-1.01)**	.29	1056	.87 (.67-1.07)	.27
logearnings 83	1230	**.89 (.77-1.00)**	.19	438	.82 (.57-1.07)	.08
logearnings 93	834	.95 (.77-1.14)	.17	219	.69 (.32-1.06)	.21
logincome 83	1306	.94 (.82-1.07)	.19	776	1.05 (.75-1.35)	.11
logincome 93	1004	**.84* (.71-.97)**	.19	551	1.19 (.80-1.57)	.16
net wealth 93	1097	.86 (73-1.00)	.14	722	1.05 (.77-1.33)	.04

4.1 References

[1] Cramer, J.S. (2011) Childhood Intelligence and Adult Mortality in the Brabant Data Set: Technical Report.

[2] Provinciaal Bestuur van Noord-Brabant (1957) *Rapport over een onderzoek naar de stand van het Gewoon Lager Onderwijs in Noord-Brabant.*

[3] Snijders, J.Th., C.J.M.H.Souren en W.J.Welten (1962) *Prestatietests Schoolvorderingen en Intelligentie.* (second edition). Groningen: Wolters.

[4] Raven, J.C. (1941) Standardisation of progressive matrices. *British Journal of Medical Psychology*, vol. 19, p.137-150.

[5] Carpenter, Patricia A., Marcel Adam Just, and Peter Shell (1990). What One Intelligence Test Measures: A Theoretical Account of the Processing in the Raven Progressive Matrices Test. *Psychological Review*, vol.97, p. 404-431.

[6] Batty, G.David, Ian J. Deary and Linda S. Gottfredson (2007) Premorbid (early life) IQ and Later Mortality Risk: Systematic Review. *Annals of Epidemiology*, vol.17, no. 4, p. 278-288.

[7] Calvin, Catherine M., Ian J. Deary, Candida Fenton, Beverly A. Roberts, Geoff Der, Nicola Leckenby and G David Batty (2010) Intelligence in youth and all-cause-mortality: systematic review with meta-analysis. *International Journal of Epidemiology*, forthcoming; doi:10.1093/ije/dyq190.

[8] Whalley, Lawrence J., and Ian J. Deary (2001) Longitudinal Cohort Study Of Childhood Iq and Survival Up To Age 76. *British Medical Journal,* vol.322, p.1-5.

[9] Hart, Carole L., Michelle D. Taylor, George Davey Smith, Lawrence J. Whalley, John M. Starr, David J. Hole, Valerie Wilson and Ian J. Deary (2003) Childhood IQ, Social Class, Deprivation, and Their Relationships with Mortality and Morbidity Risk in Later Life: Prospective Observational Study Linking the Scottish Mental Survey 1932 and the Midspan studies. *Psychosomatic Medicine*, vol. 65, p.877-883.

[10] Osler, M., A-M. N. Andersen, P. Due, R. Lund, M.T. Damsgaard and B.E. Holstein (2003) Socioeconomic position in early life, birth weight, childhood cognitive function, and adult mortality. A longitudinal study of Danish men born in 1953. *Journal of Epidemiology and Community Health*, vol.57, nr 9, p.681-686.

[11] Pearce, M.S., I.J.Deary, A.H. Young & L. Parker (2006) Childhood IQ and deaths up to middle age: the Newcastle Thousand Families Study. *Public Health*, vol.120, p..1020-26.

[12] Deary, Ian J., G.David Batty, Alison Pattie and Catharine R. Gale (2008) More Intelligent, More Dependable Children Live Longer: A 55-Year Longitudinal Study of a Representative Sample of the Scottish Nation. *Psychological Science*, vol.19, p. 874-880.

[13] Jokela, Markus, G.David Batty, Ian J. Deary. Catharine R. Gale & Mika Kivimaki (2009) Low childhood IQ and adult mortality: the role of explanatory factors in the 1958 British birth cohort. *Pediatrics*, vol.124, p.e380-388.

[14] Lager, A., S. Bremberg & D. Vagero. The association of early IQ and education with mortality: 65 year longitudinal study in Malmo, Sweden. *British Medical journal*, vol.339, b5282 Doi:10.1136/bmj.b5282

[15] Deary, Ian (2008) Why do intelligent people live longer? *Nature*, vol. 456, p.175-176.

[16] Deary, I.J., L.J.Whalley, H. Lemon, J.R.Crawford and J.M. Starr (2000) The Stability of Individual Differences in Mental Ability from Childhood to Old Age: Follow-up of the 1932 Scottish Mental Survey. *Intelligence*, Vol.28, p.49-55.

4.2 Appendix. ML Estimates

This appendix gives the Maximum Likelihood estimates of equation (1), with the effective sample size n and the number of dead m. We also report the standard deviation of the explanatory variable X since it enters into the calculation of HRs. - The value of the mean sample loglikelihood varies with the extent of the right-hand censoring (see Figure 6 of [1]), and this explains why logL/n is so much higher for women than for men.

Table A1 Estimates for Tables 1 and 3

variable	n	m	β_1	s.d.β_1	s.d.X	logL/n
			1790	MEN		
social class	1622	225	-.1073	.0974	.6807	-1.5850
working child	1575	221	-.0814	.2345	.2936	-1.6023
# children	1642	230	-.0467	.0246	2.7653	-1.6000
birthrank	1635	228	-.0252	.0268	2.5733	-1.5937
education F	1268	176	.0762	.0858	.8275	-1.5876
education M	1253	177	-.0717	.1294	.6288	-1.6022
logIQ	1600	222	-1.3155	.4795	.1400	-1.5843
logIQa	1463	202	-1.4320	.5074	.1408	-1.5770
logVOC	1456	200	-.3078	.5130	.1376	-1.5721
teacher	1668	232	-.0499	.0512	1.3433	-1.5904
forms rep'd	1690	235	.0009	.0935	.6822	-1.5899
			1208	WOMEN		
social class	1095	93	.1345	.1591	.6622	-1.0157
working child	1058	93	.2183	.2747	.3490	-1.0456
# children	1179	100	-.0046	.0360	2.7782	-1.0203
birthrank	1172	100	.0292	.0383	2.5237	-1.0185
education F	680	50	-.2831	.2067	.8734	-.8913
education M	682	50	-.1100	.2540	.6144	-.8904
logIQ	1132	95	-1.4013	.7364	.1403	-1.0003
logIQa	1125	91	-1.8276	.7553	.1392	-.9658
logVOC	1121	91	-1.0840	.7551	.1389	-.9707
teacher	1161	100	-.0503	.0876	1.1717	-1.0268
forms rep'd	1178	101	.1165	.1544	.6021	-1.0229

Table A 2 Estimates for Table 4

variable	n	m	β_1	s.d.β_1	s.d.X	logL/n
			1709 MEN			
education	1589	207	-.1377	.0792.	.9597	-1.4961
log earnings 83	1230	164	-.2087	.1188.	.5653	-1.5268
log earnings 93	834	100	-.0805	.1582	.6262	-1.3918
log family income 83	1306	179	-.1068	.1222	.5571	-1.5660
log family income 93	1004	136	-.2893	.1235	.5902	-1.5524
net wealth 93	1097	137	-.0602	.0320	2.4911.	-1.4422
			1208 WOMEN			
education	1056	91	-.1729	.1464.	.7961	–1.0284
log earnings 83	438	32	-.1719	.1340.	1.1568	–.8845
log earnings 93	219	14	-.4153	.3056.	.8993	–.7721
log family income 83	776	54	.0560	.1627.	.8859	–.8457
log family income 93	551	43	.2390	.2300.	.7124	–.9368
net wealth 93	722	56	.0192	.0561.	2.4330	–.9322

Table 3 gives the estimates for two variables, logIQa (with coefficient β_1) and some other variable X_2 (with coefficient β_2). These are referred to in the text of Section 3.2.

Table A3 Estimates for two explanatory variables

X_2	n	m	β_1	s.d.β_1	β_2	s.d. β_2	logL/n
			1790 MEN				
log IQ	1439	197	-1.1115	.5286	-.6997	.6321	-1.5729
forms rep'd	1400	194	-1.4574	.5231	.0330	.1030	-1.5817
social class	1351	189	-1.4708	.5301	-.0210	.1068	-1.5950
working child	1290	180	-1.6429	.5371	-.1656	.2498	-1.5901
# of children	1169	191	-1.6599	.5299	-.0678	.0273	-1.5899
birthrank	1360	189	-1.6387	.5322	-.0425	.0300	-1.5849
education F	1058	143	-1.4591	.6238	.0856	.1066	-1.5483
education M	1054	144	-1.3989	.6145	-.0128	.1456	-1.5641
			1208 WOMEN				
log IQ	1101	91	-1.5259	.9181	-.3598	.9029	-.9853
forms rep'd	1097	90	-1.7102	.7786	.1752	.1632	-.9778
social class	1018	83	-2.5747	.8042	.0965	.1703	-.9717
working child	979	82	-1.8002	.8028	.2924	.2791	-.9954
# children	1090	89	-1.7170	.7743	.0057	.0382	-.9745
birthrank	1092	89	-1.6992	.7742	.0377	.0400	-.9724
education F	626	41	-1.2266	1.2358	-.3190	.2545	-.7980
education M	628	41	-1.5151	1.1918	-.1164	.2916	-.7970