**CHDS MEDCIAL CONTACTS – BIRTH TO AGE 13**

**BACKGROUND**

This data dictionary refers to chds20110906.sas7bdat.

File chds20110906.sas7bdat is a SAS data set that has combined:

* the original chds05.sas7bdat dataset (with data on children from years 0 to 5),
* chds10.sas7bdat (data sent by JH in July 11 including Burt reading score, conduct problems, mother’s emotional responsiveness and punitiveness and preschool education) and
* chds613extravars.sas7bdat (data sent by JH in September 11 including data on health service use and all intermediate variables in chds05.sas7bdat for years 6 to 13).

Note that for chds613extravars.sas7bdat (and hence for chds20110906.sas7bdat) we have intermediate variables up to age 13 but health service use only up to age 10.

The original chds05.sas7bdat dataset contained information on 169 variables describing the history of medical contacts in the CHDS cohort over the period from birth to 5 years, together with a series of potentially relevant predictors describing social, family and related factors. It contained data for all 1265 children who entered the study at birth. Each case is identified by a unique code number (a0). The data were stored as a single record per case.

chds20110906.sas7bdat is in ‘tall’ format to aid in analysis with SAS– i.e the data are stored as 13 records per person. The original yearly variables (i.e with suffixes 1 to 5) are still included, and repeated within each child’s record; however, new un-suffixed variables (some with a new prefix of ‘z1’ - see below) containing the same information in a ‘tall’ fashion (i.e corresponding to the newly included ‘**age’** of child variable) have been created. Please note some exceptions to the ‘un-suffixed’ convention –the original variables **fage1, fage2, fage3, fage4, and fage5**, have been put into tall form in the **fage\_years** variable, because **fage** was already in the original set as the father’s age at the birth of the child. The same applies for the ‘mage’ group of variables**.**

New variables giving the value of a variable in the previous year have also been created and are stored in variables suffixed as ‘\_previous’.

The variables have sometimes been coded in a way perhaps better suited to analysis in SAS. When this has happened a prefix of ‘z1’ or ‘AGG’ has been added: those with the ‘z1’ prefix show that it is the original variable coded as ‘0 /1’ with 0 being the desired reference group for statistical analysis; those with the prefix ‘AGG’ show that it is the original variable aggregated up, to produce a better variable for analysis. Also - the ‘**r1stchildethn**’ gives a child’s ethnicity based on their parent’s ethnicity, in such a way that the desired reference group for analysis is coded as a ‘1’. ‘**r1stfethn**’ and ‘**r1stmethn**’ give the original parent ethnicities but again coded so that the desired reference group for analysis is now ‘1’. ‘**r1stfeduc**’ and ‘**r1stmeduc**’ do the same for parent education levels.

All reference groupings are to allow for use with ‘param=ref ref=first’ coding in proc logistic.

Details of all new variables are provided in italics, alongside the details for the original variables.

The medical data comprise counts of the number of medical contacts in each year from birth to age 5 for a limited range of conditions. Specifically, in the first instance John provided data on contacts for respiratory illness and accidental injury (the two most common reasons for medical contacts in childhood) as well as for total morbidity contacts in each year. The data are supplied separately for GP/doctor visits, hospital admissions and hospital outpatient attendances. He has also provided data on doctor contacts for preventive health care (immunizations, well-child checks, etc). I am happy to consider other classes of morbidity at a later date, but I have gone for simplicity in the first instance.

The table below summarises the number of children assessed at each age from birth to age 5. Because the sample sizes vary with age, the numbers with non-missing data on medical contacts (and other variables) for each year also vary with age. In addition, for some variables there is a small amount of additional missing data due to reasons other than the child not being assessed. Thus, for example, there were 1180 children assessed at age 1 year, but the medical contact variables for the period birth to age 1 were assessed on samples ranging from 1171 to 1177 cases.

Age Number Assessed

Birth 1262

1 year 1180

2 years 1156

3 years 1143

4 years 1127

5 years 1123

For year 6 to 13, Jessica is not sure how Janet got the above table so it is not provided at this time.

The following pages document the variables in more detail. Please note that there are coding conventions for some variables that will need to be taken into account in any analysis eg how we code father’s age when then is no father in the family at a given age. These coding conventions are described below.

**MEDICAL CONTACTS 0-13 YEARS**

**General Practitioner/Medical Specialist (non hospital) Contacts**

The following variables provide measures of the number of (non-hospital) doctor contacts for a range of conditions in each 12 month period from birth to age 7 years. The majority of these contacts are with family GPs, but contacts may also include visits to private medical specialists.

Please note: These variables will produce slightly different rates of contact to those reported in our paper on child morbidity in the preschool years (Aust Paed J 1985, 21: 139-145). The counts have been calculated in a slightly different way from the original paper.

*Age child ‘s age: 1-5*

*Age\_minus1 child ‘s age minus one: 0-4*

gpprev1 GP visits preventive health care 0-1 yr

gpprev2 GP visits preventive health care 1-2 yr

gpprev3 GP visits preventive health care 2-3 yr

gpprev4 GP visits preventive health care 3-4 yr

gpprev5 GP visits preventive health care 4-5 yr

gpprev6 GP visits preventive health care 5-6 yr

gpprev7 GP visits preventive health care 6-7 yr

*gpprev GP visits preventive health care in year as given by the* **age** *variable*

*gpprev\_previous GP visits preventive health care from previous year (i.e from when survey child was* **age\_minus1***)*

The number of general practitioner/doctor visits for preventive health care for each 12 month period from birth to age 10 years. Preventive health care includes visits for immunisations, postnatal and well-child checks.

gpresp1 GP visits respiratory illness 0-1 yr

gpresp2 GP visits respiratory illness 1-2 yr

gpresp3 GP visits respiratory illness 2-3 yr

gpresp4 GP visits respiratory illness 3-4 yr

gpresp5 GP visits respiratory illness 4-5 yr

...

gpresp10 GP visits respiratory illness 9-10 yr

*gpresp GP visits respiratory illness in year as given by the* **age** *variable*

*gpresp\_previous GP visits respiratory illness from previous year (i.e from when survey child was* **age\_minus1***)*

The number of general practitioner/doctors visits for respiratory illness for each 12 month period from birth to 10 years. Respiratory illness includes all contacts for bronchitis, pneumonia, colds, coughs, flu, sore throats, ear infection, asthma, and other miscellaneous respiratory conditions. These counts are limited to a maximum of 12 contacts within a year.

gpacc1 GP visits accidental injury 0-1 yr

gpacc2 GP visits accidental injury 1-2 yr

gpacc3 GP visits accidental injury 2-3 yr

gpacc4 GP visits accidental injury 3-4 yr

gpacc5 GP visits accidental injury 3-5 yr

...

gpacc10 GP visits accidental injury 9-10 yr

I suspect they made a mistake for gpacc5. I think it should be “4-5 yr” not “3-5 yr”.

*gpacc* *GP visits accidental injury in year as given by* **age** *variable*

*gpacc\_previous* *GP visits accidental injury in previous year (i.e from when survey child was* **age\_minus1***)*

The number of general practitioner/doctors visits for accidental injuries for each 12 month period from birth to 7 years. Includes all contacts for accidental poisoning, burns/scalds, fractures, cuts, abrasion, contusions, concussions, sprains, etc. These counts are limited to a maximum of 12 contacts within a year.

gpmorb1 GP total morbidity visits 0-1 yr

gpmorb2 GP total morbidity visits 1-2 yr

gpmorb3 GP total morbidity visits 2-3 yr

gpmorb4 GP total morbidity visits 3-4 yr

gpmorb5 GP total morbidity visits 4-5 yr

gpmorb6 GP total morbidity visits 5-6 yr

gpmorb7 GP total morbidity visits 6-7 yr

*gpmorb* *GP total morbidity visits in year as given by* **age** *variable*

*gpmorb\_previous* *GP total morbidity visits in previous year (i.e from when survey child was* **age\_minus1***)*

The number of general practitioner/doctors visits for any form of morbidity for each 12 month period from birth to 10 years. This includes all contacts for a physical or mental health problem in a given year (including the visits for respiratory illness and/or accidents above) These counts are limited to a maximum of 12 contacts within a year.

gptotvis1 GP total consultations 0-1 yr

gptotvis2 GP total consultations 1-2 yr

gptotvis3 GP total consultations 2-3 yr

gptotvis4 GP total consultations 3-4 yr

gptotvis5 GP total consultations 4-5 yr

...

gptotvis10 GP total consultations 9-10 yr

*gptotvis* *GP total consultations in year as given by* **age** *variable*

*gptotvis\_previous* *GP total consultations in previous year (i.e from when survey child was* **age\_minus1***)*

The total number of general practitioner/doctors visits for any reason (including all morbidity contacts and all preventive health care contacts) for each 12 month period from birth to age 5 years. These variables have not been truncated to an upper limit.

**Hospital Admissions**

The following variables provide measures of the number of hospital admissions for a range of conditions in each 12 month period from birth to age 10 years.

Please note: These variables will produce slightly different numbers to those reported in our paper on child morbidity in the preschool years (Aust Paed J 1985, 21: 139-145). The counts have been calculated in a slightly different way from the original paper.

hadmresp1 Hosp adms respiratory 0-1 yr

hadmresp2 Hosp adms respiratory 1-2 yr

hadmresp3 Hosp adms respiratory 2-3 yr

hadmresp4 Hosp adms respiratory 3-4 yr

hadmresp5 Hosp adms respiratory 4-5 yr

...

hadmresp10 Hosp adms respiratory 9-10 yr

The number of hospital admissions for respiratory illness for each 12 month period from birth to age 10 years. Respiratory illness defined as above.

*hadmresp Hosp adms respiratory in year as given by* **age** *variable*

*hadmresp \_previous Hosp adms respiratory in previous year (i.e from when survey child was* **age\_minus1** *variable)*

hadmacc1 Hosp adms accid injury 0-1 yr

hadmacc2 Hosp adms accid injury 1-2 yr

hadmacc3 Hosp adms accid injury 2-3 yr

hadmacc4 Hosp adms accid injury 3-4 yr

hadmacc5 Hosp adms accid injury 4-5 yr

...

hadmacc10 Hosp adms accid injury 9-10 yr

*hadmacc Hosp adms accid injury in year as given by* **age** *variable*

*hadmacc \_previous Hosp adms accid injury in previous year (i.e from when survey child was* **age\_minus1***)*

The number of hospital admissions for accidental injury for each 12 month period from birth to age 10 years. Accidental injury defined as above.

hadmtot1 Total hospital admissions 0-1 yr

hadmtot2 Total hospital admissions 1-2 yr

hadmtot3 Total hospital admissions 2-3 yr

hadmtot4 Total hospital admissions 3-4 yr

hadmtot5 Total hospital admissions 4-5 yr

...

hadmtot10 Total hospital admissions 9-10 yr

*hadmtot Total hospital admissions in a year as given by* **age** *variable*

*hadmtot \_previous Total hospital admissions in previous year (i.e from when survey child was* **age\_minus1** *variable)*

The total number of hospital admissions for any reason (including respiratory and accident admissions as above) for each 12 month period from birth to age 10 years. Includes both arranged and acute admissions.

hnights1 Total nights in hospital 0-1 yr

hnights2 Total nights in hospital 1-2 yr

hnights3 Total nights in hospital 2-3 yr

hnights4 Total nights in hospital 3-4 yr

hnights5 Total nights in hospital 4-5 yr

...

hnights10 Total nights in hospital 9-10 yr

*hnights Total nights in hospital in a year as given by* **age** *variable*

*hnights \_previous Total nights in hospital in previous year (i.e from when survey child was* **age\_minus1***)*

The total number of nights spent in hospital as a result of hospital admission for each 12 month period from birth to age 10 years.

**Hospital Outpatient Attendances**

The following variables provide measures of the number of hospital outpatient attendances for a range of conditions in each 12 month period from birth to age 10 years. Outpatient attendances include both acute visits (eg to the Emergency Department) and arranged visits for check-ups, follow-ups or investigation of conditions.

Please note: These variables will produce slightly different numbers to those reported in our paper on child morbidity in the preschool years (Aust Paed J 1985, 21: 139-145). The counts have been calculated in a slightly different way from the original paper.

houtpresp1 Hosp outpat respiratory 0-1 yr

houtpresp2 Hosp outpat respiratory 1-2 yr

houtpresp3 Hosp outpat respiratory 2-3 yr

houtpresp4 Hosp outpat respiratory 3-4 yr

houtpresp5 Hosp outpat respiratory 4-5 yr

...

houtpresp10 Hosp outpat respiratory 9-10 yr

*houtpresp Hosp outpat respiratory in year as given by* **age** *variable*

*houtpresp\_previous Hosp outpat respiratory injury in previous year (i.e from when survey child was* **age\_minus1***)*

The number of hospital outpatient attendances for respiratory illness for each 12 month period from birth to age 10 years. Respiratory illness defined as above.

houtpacc1 Hosp outpat accid injury 0-1 yr

houtpacc2 Hosp outpat accid injury 1-2 yr

houtpacc3 Hosp outpat accid injury 2-3 yr

houtpacc4 Hosp outpat accid injury 3-4 yr

houtpacc5 Hosp outpat accid injury 4-5 yr

...

houtpacc10 Hosp outpat accid injury 9-10 yr

*houtpacc Hosp outpat accid injury* *in year as given by* **age** *variable*

*houtpacc \_previous Hosp outpat accid injury in previous year (i.e from when survey child was* **age\_minus1***)*

The number of hospital outpatient attendances for accidental injury for each 12 month period from birth to age 10 years. Accidental injury defined as above.

houtptot1 Hosp outpat total visits 0-1 yr

houtptot2 Hosp outpat total visits 1-2 yr

houtptot3 Hosp outpat total visits 2-3 yr

houtptot4 Hosp outpat total visits 3-4 yr

houtptot5 Hosp outpat total visits 4-5 yr

...

houtptot10 Hosp outpat total visits 9-10 yr

*houtptot Hosp outpat total visits* *in year as given by* **age** *variable*

*houtptot \_previous Hosp outpat total visits in previous year (i.e from when survey child was* **age\_minus1***)*

The number of hospital outpatient attendances for any reason (including respiratory and accident visits as above) for each 12 month period from birth to age 10 years.

**SOCIAL, FAMILY AND OTHER FACTORS**

The following variables describe a range of social, family and individual characteristics/ circumstances assessed over the period from birth to age 13 years. Wherever possible the measures have been replicated for each year, to allow for changing family circumstances over time.

**Code Number**

a0 The four digit code number assigned to each child

**Child Characteristics**

gender Child’s sex, coded as:

1 = Male

2 = Female

*z1gender Child’s sex, coded as:*

*0=Female*

*1=Male*

bthorder Birth order in the family (range 1-5+)

*z1bthorder Child’s bthorder, coded as:*

*0=1st*

*1=2nd or above*

breast Duration of breast feeding (months). Range 0-12, a code of 12 means 12 months or longer.

*z1breast Number of months breastfed, coded as:*

*0=One month or more*

*1=Not breastfed*

twin Twin birth, coded as:

1 = Singleton

2 = Twin

*z1twin Twin birth coded as:*

*0=Singleton*

*1=Twin*

*r1stchildethn Child ethnicity, based on mother and father’s ethnicities at birth of child (using the prioritisation order as follows: Maori, then pacific, then Other).*

*Coded on a 3 point scale (with reference for SAS analysis being ‘1’ or ‘Other’) as follows:*

*3 = Maori/part Maori*

*2 = Pacific Island*

*1 = Other*

**Antenatal/Perinatal Factors**

bw Birthweight (grams). Range 1100-5140.

*z1bw Birthweight coded as:*

*0= 3360 grams, or greater*

*1=less than 3360 grams*

*bwkg* birthweight (kilograms)

*bwkg = bw/1000*

ga Gestational age at birth – weeks. Range 28-44

*z1ga Gestational age at birth coded as:*

*0= 40 weeks, or greater*

*1= less than 40 weeks*

nicu Child admitted to neonatal intensive care unit at birth, coded as:

1 = Child admitted to NICU

2 = Child not admitted to NICU

*z1nicu 0=* Child not admitted to NICU

*1=* Child admitted to NICU

pregsmk Average number of cigarettes smoked per day by natural mother during pregnancy. Range 0-50.

*z1pregsmoke* **pregsmk** *coded as:*

*0=mother didn’t smoke in pregnancy*

*1=mother smoked in pregnancy*

*catpregsmk* **pregsmk** *c*oded as:

1 if pregsmk = 0

2 if 1 <= pregsmk <= 5 (between 1 and 5 inclusive)

3 if 6 <= pregsmk <= 10

4 if 11 <= pregsmk <= 20

5 if 21 <= pregsmk <= 50 (50 is the maximum value)

*catpregsmk2* the same as *catpregsmk* except the value of this variable is the median of pregsmk for each catpregsmk group. E.g. in catpregsmk group 2 (1 to 5 cigarettes per day) the median number of cigarettes smoked per day is 3. This variable was coded this way so it could be treated as a continuous variable in the model if desired (it keeps accurate distances between the categories).

So *catpregsmk2* is **pregsmk** coded as:

0 if pregsmk = 0

3 if 1 <= pregsmk <= 5 (between 1 and 5 inclusive)

8 if 6 <= pregsmk <= 10

16 if 11 <= pregsmk <= 20

27 if 21 <= pregsmk <= 50 (50 is the maximum value)

*catpregsmk3* **pregsmk** coded as:

1. if pregsmk = 0
2. if 1 <= pregsmk <= 10
3. if pregsmk >= 11

catpregsmk4 A categorisation used by University of Otago Christchurch (previously called Christchurch School of Medicine)

**pregsmk** coded as

1. if pregsmk = 0
2. if 1 <= pregsmk <= 19
3. if pregsmk >= 20

pregalc Average number of alcoholic drinks per week consumed by natural mother during pregnancy. Range 0-61.

*z1pregalc* **pregalc** *coded as:*

*0=mother didn’t drink alcohol in pregnancy*

*1=mother drank alcohol in pregnancy*

*pregalc2* **pregalc** with values greater than 35 set to 25.

There were two outlying values for pregalc. Child 348 had a mother who drank 36 drinks a week, on average, during pregnancy and child 881 had a mother who drank 61 drinks a week, on average, during pregnancy. These were set to be 25 so as to enable better modelling.

*catpregalc* **pregal**c coded as:

1. if pregalc = 0
2. if 1 <= pregalc <= 7 (between 1 and 7 inclusive)
3. if pregalc > 7

*catpregalc2* **pregalc** coded as:

1. if pregalc = 0
2. if 1 <= pregalc <= 6
3. if pregalc >= 7

**Parental Characteristics**

MAGE Maternal age in years at birth of survey child. In continuous form.

*z1mage* **mage** *coded as:*

*0= 25 years or greater*

*1= less than 25 years*

mage1 Maternal age when survey child aged 1 yr

mage2 Maternal age when survey child aged 2 yrs

mage3 Maternal age when survey child aged 3 yrs

mage4 Maternal age when survey child aged 4 yrs

mage5 Maternal age when survey child aged 5 yrs

*mage\_years* *Mothers age when survey child age =* **age** *variable.* Assessed in whole years. A code of 99 implies there was no mother (figure) in the home i.e. the family was a solo father family at the time of the interview.

*mage\_years\_previous* *Mothers age in previous year* *(i.e. when survey child was aged* **age\_minus1***)*

We do not have mage\_years for years 6 to 13 but it would be good to get it. – Need to check whether this statement is still true – we may have got it now.

*no99mage1 Maternal age when survey child aged 1 yr* ***,*** *as above but 99’s made missing.*

*no99mage2 Maternal age when survey child aged 2 yrs****,*** *as above but 99’s made missing.*

*no99mage3 Maternal age when survey child aged 3 yrs****,*** *as above but 99’s made missing.*

*no99mage4 Maternal age when survey child aged 4 yrs****,*** *as above but 99’s made missing.*

*no99mage5 Maternal age when survey child aged 5 yrs****,*** *as above but 99’s made missing.*

*no99mage\_years* *Mothers age when survey child age =* **age** *variable***,** *as above but 99’s made missing.*

*no99mage\_years\_previous* *Mothers age in previous year* *(i.e when survey child was aged* **age\_minus1***)* **,** *as above but 99’s made missing.*

fage Father’s age at birth of survey child. Coded as:

1 = <20 years

2 = 20-24 years

3 = 25-29 years

4 = 30-34 years

5 = 35-39 years

6 = 40+ years

Information is recorded on father’s age at the child’s birth even if the father was not resident with the child.

*z1fage* **fage** *coded as:*

*0=30 years or older*

*1=29 years or younger*

fage1 Father’s age when survey child aged 1 yr

fage2 Father’s age when survey child aged 2 yrs

fage3 Father’s age when survey child aged 3 yrs

fage4 Father’s age when survey child aged 4 yrs

fage5 Father’s age when survey child aged 5 yrs

Father’s age assessed in whole years. A code of 99 implies there was no father (figure) in the home i.e. the family was a solo mother family at the time of the interview.

*fage\_years Father’s age when survey child was aged* **age** *variable. Missing if ‘99’.* We do not have fage for years 6 to 13 but it would be good to get it.

*fage\_years\_previous Father’s age from previous year (i.e from when survey child was aged* **age\_minus1***). Missing if ‘99’.*

*no99fage1 Paternal age when survey child aged 1 yr,* ***,*** *as above but 99’s made missing.*

*no99fage2 Paternal age when survey child aged 2 yrs****,*** *as above but 99’s made missing.*

*no99fage3 Paternal age when survey child aged 3 yrs****,*** *as above but 99’s made missing.*

*no99fage4 Paternal age when survey child aged 4 yrs****,*** *as above but 99’s made missing.*

*no99fage5 Paternal age when survey child aged 5 yrs****,*** *as above but 99’s made missing.*

*no99fage\_years* *Fathers age when survey child age =* **age** *variable***,** *as above but 99’s made missing.*

*no99fage\_years\_previous* *Fathers age in previous year* *(i.e when survey child was aged* **age\_minus1***)* **,** *as above but 99’s made missing.*

methn Maternal ethnicity at birth of survey child

fethn Paternal ethnicity at birth of survey child

These variables are coded on a 3 point scale as follows:

1 = Maori/part Maori

2 = Pacific Island

3 = Other

Note that information is recorded on paternal ethnicity at the point of birth (where known) even if the father was not resident with the child.

*r1stmethn*

*r1stfethn*

*These variables coded on a 3 point scale (with reference for SAS analysis being ‘1’ or ‘Other’) as follows:*

*3 = Maori/part Maori*

*2 = Pacific Island*

*1 = Other*

meduc Maternal education at survey child’s birth

feduc Paternal education at survey child’s birth

These variables are coded on a three point scale as follows:

1 = No formal educational qualifications

2 = Secondary qualifications (School Cert, University Entrance, Bursary, etc.)

3 = Tertiary qualifications (University degree, tertiary technical diploma, etc.)

Note that information is recorded on paternal education at the point of birth even if the father was not resident with the child.

*r1stmeduc*

*r1stfeduc*

*These variables coded on a 3 point scale (with reference for SAS analysis being ‘1’ or ‘Tertiary qualifications’) as follows:*

*3 =* No formal educational qualifications

*2 =* Secondary qualifications (School Cert, University Entrance, Bursary, etc.)

*1 =* Tertiary qualifications (University degree, tertiary technical diploma, etc.)

msmoke1 maternal smoking (cigs/day) 1 yr

msmoke2 maternal smoking (cigs/day) 2 yr

msmoke3 maternal smoking (cigs/day) 3 yr

msmoke4 maternal smoking (cigs/day) 4 yr

msmoke5 maternal smoking (cigs/day) 5 yr

…

msmoke13 maternal smoking (cigs/day) 13 yr

Note: Where there is no mother figure in the house in a given year (ie single father family) these variables have been coded to zero.

*msmoke maternal smoking (cigs/day)- year=* **age** *variable*

*msmoke\_previous maternal smoking (cigs/day)- year=* **age\_minus1** *variable*

*z1msmoke1 0=mother doesn’t smoke, 1= mother smokes 1 yr*

*z1msmoke2 0= mother doesn’t smoke, 1= mother smokes 2 yr*

*z1msmoke3 0= mother doesn’t smoke, 1= mother smokes 3 yr*

*z1msmoke4 0= mother doesn’t smoke, 1= mother smokes 4 yr*

*z1msmoke5 0= mother doesn’t smoke, 1= mother smokes 5 yr*

*…*

*z1msmoke13 0= mother doesn’t smoke, 1= mother smokes 13 yr*

*z1msmoke 0= mother doesn’t smoke, 1= mother smokes during year=* **age** *variable*

*z1msmoke\_previous mother smoking status in previous year: 0=didn’t smoke, 1=smoked during year=* **age\_minus1** *variable*

fsmoke1 Paternal smoking (cigs/day) 1 yr

fsmoke2 Paternal smoking (cigs/day) 2 yr

fsmoke3 Paternal smoking (cigs/day) 3 yr

fsmoke4 Paternal smoking (cigs/day) 4 yr

fsmoke5 Paternal smoking (cigs/day) 5 yr

…

fsmoke13 Paternal smoking (cigs/day) 13 yr

Note: Where there is no father figure in the house in a given year (ie single mother family) these variables have been coded to zero.

*fsmoke paternal smoking (cigs/day)- year=* **age** *variable*

*fsmoke\_previous paternal smoking (cigs/day)- year=* **age\_minus1** *variable*

*z1fsmoke1 0=father doesn’t smoke, 1= father smokes 1 yr*

*z1fsmoke2 0=father doesn’t smoke, 1= father smokes 2 yr*

*z1fsmoke3 0=father doesn’t smoke, 1= father smokes 3 yr*

*z1fsmoke4 0=father doesn’t smoke, 1= father smokes 4 yr*

*z1fsmoke5 0=father doesn’t smoke, 1= father smokes 5 yr*

*…*

*z1fsmoke13 0=father doesn’t smoke, 1= father smokes 13 yr*

*z1fsmoke 0=father doesn’t smoke, 1= father smokes during year =* **age** *variable*

*z1fsmoke\_previous 0=father didn’t smoke, 1= father smoked during year =* **age\_minus1** *variable*

**Family Socioeconomic Circumstances**

SESBTH Family socioeconomic status at birth of survey child

Based on Elley-Irving (1976) classification of male occupational status for NZ, coded into 3 levels as follows:

1 = Professional, managerial

2 = Clerical, technical, skilled

3 = Semi-skilled, unskilled, unemployed

grossinc1 Gross weekly family income ($) at 1 yr (1978)

grossinc2 Gross weekly family income ($) at 2 yr (1979)

grossinc3 Gross weekly family income ($) at 3 yr (1980)

grossinc4 Gross weekly family income ($) at 4 yr (1981)

grossinc5 Gross weekly family income ($) at 5 yr (1982)

We did not get income data for years 6 to 13.

*grossinc Gross weekly family income ($) when survey child aged* **age** *variable*

*grossinc\_previous Gross weekly family income ($) in previous year (i.e from when survey child was aged* **age\_minus1***)*

Note: There is quite a substantial amount of missing data on family incomes (10% approximately in any year) due to families being reluctant to divulge income levels. In some years a few families have a gross income estimate of zero: these families were usually either dependent on someone else for financial support or were in some form of transition e.g. living off savings between jobs. At the other end of the spectrum, in some years the upper level of our family income codes was set too low and this resulted in a substantial minority of families (up to 10%) falling into the highest value category for gross family income. Finally, if you need to calculate income estimates adjusted for inflation I would suggest that you use the June quarter annual adjustments, since our interview period was from April to August each year.

avinc5 Average family income decile (0-5 years)

To avoid difficulties with missing data and inflation adjustment of income data you may prefer to use this variable. It is calculated by first classifying families into deciles of family income for each year and then averaging the resulting measures over all years with non-missing data to derive an average family income decile for the period 0-5 years. It is thus a global measure of the level of income available to the family over the 5 year period. The decile rank has been multiplied by 10. Thus the variable ranges from 10-100, with a score of 10 implying family consistently in the lowest decile of income, and a score of 100 implying family consistently in the highest decile of income over the period.

welfare1 Family in receipt of welfare benefits at 1 yr

welfare2 Family in receipt of welfare benefits at 2 yr

welfare3 Family in receipt of welfare benefits at 3 yr

welfare4 Family in receipt of welfare benefits at 4 yr

welfare5 Family in receipt of welfare benefits at 5 yr

…

welfare13 Family in receipt of welfare benefits at 6 yr

*welfare Family in receipt of welfare benefits when survey child aged =* **age** *variable*

*welfare\_previous Family in receipt of welfare benefits in previous year (i.e from when survey child was* **age\_minus1***)*

These are dichotomous variables, coded as:

0 = Family not in receipt of welfare benefit

1 = Family in receipt of welfare benefit

In the great majority of cases the welfare benefit received was the Domestic Purposes Benefit. The remaining benefits were typically sickness benefits or the disabled child allowance. Unemployment was not an issue for cohort families at this point in time.

mhrswrk1 Maternal hours per week paid employment at 1 yr

mhrswrk2 Maternal hours per week paid employment at 2 yr

mhrswrk3 Maternal hours per week paid employment at 3 yr

mhrswrk4 Maternal hours per week paid employment at 4 yr

mhrswrk5 Maternal hours per week paid employment at 5 yr

…

mhrswrk13 Maternal hours per week paid employment at 6 yr

If mother not working or there was no mother present in the household in a given year then the corresponding measure has been coded as zero. A code of 97 or 98 should be taken to mean 97 or more hours per week.

*mhrswrk Maternal hours per week paid employment -when child aged* ***age*** *variable*

*mhrswrk\_previous Maternal hours per week paid employment- in previous year (i.e from when child was aged* ***age\_minus1*** *variable)*

*z1mhrswrk1 0=mother has paid employment , 1=mother 0 hours paid employment -yr1*

*z1mhrswrk2 0=mother has paid employment , 1=mother 0 hours paid employment -yr2*

*z1mhrswrk3 0=mother has paid employment , 1=mother 0 hours paid employment -yr3*

*z1mhrswrk4 0=mother has paid employment , 1=mother 0 hours paid employment -yr4*

*z1mhrswrk5 0=mother has paid employment, 1=mother 0 hours paid employment -yr5*

*…*

*z1mhrswrk13 0=mother has paid employment, 1=mother 0 hours paid employment –yr13*

*z1mhrswrk 0=mother has paid employment, 1=mother 0 hours paid employment -when child aged* **age** *variable*

*z1mhrswrk\_previous 0=mother has paid employment, 1=mother 0 hours paid employment employment in previous year (i.e from when child was aged* **age\_minus1** *variable)*

fhrswrk1 Paternal hours per week paid employment at 1 yr

fhrswrk2 Paternal hours per week paid employment at 2 yr

fhrswrk3 Paternal hours per week paid employment at 3 yr

fhrswrk4 Paternal hours per week paid employment at 4 yr

fhrswrk5 Paternal hours per week paid employment at 5 yr

…

fhrswrk13 Paternal hours per week paid employment at 13 yr

If there was no father present in the household in a given year then the corresponding measure has been coded as zero. A code of 97 or 98 should be taken to mean 97 or more hours per week.

*fhrswrk paternal hours per week paid employment -when child aged* ***age*** *variable*

*fhrswrk\_previous paternal hours per week paid employment- in previous year (i.e from when child was aged* ***age\_minus1*** *variable)*

*z1fhrswrk1 0=father has paid employment , 1=father 0 hours paid employment -yr1*

*z1fhrswrk2 0=father has paid employment , 1=father 0 hours paid employment -yr2*

*z1fhrswrk3 0=father has paid employment , 1=father 0 hours paid employment -yr3*

*z1fhrswrk4 0=father has paid employment , 1=father 0 hours paid employment -yr4*

*z1fhrswrk5 0=father has paid employment, 1=father 0 hours paid employment -yr5*

*…*

*z1fhrswrk13 0=father has paid employment, 1=father 0 hours paid employment –yr13*

*z1fhrswrk 0=father has paid employment, 1=father 0 hours paid employment -when child aged* **age** *variable*

*z1fhrswrk\_previous 0=father has paid employment, 1=father 0 hours paid employment employment in previous year (i.e from when child was aged* **age\_minus1** *variable)*

sol1 Interviewer rating of family standard of living 1 yr

sol2 Interviewer rating of family standard of living 2 yr

sol3 Interviewer rating of family standard of living 3 yr

sol4 Interviewer rating of family standard of living 4 yr

sol5 Interviewer rating of family standard of living 5 yr

We did not get sol for years 6 to 13

*sol* *Interviewer rating of family standard of living at year =* **age** *variable*

*sol\_previous* *Interviewer rating of family standard of living* *from the previous year (i.e at year =* **age\_minus1**)

Interviewer ratings were made of a 5 point scale as follows:

1 = family obviously affluent

2 = above average

3 = average

4 = below average

5 = family obviously poor

*AGGsol Aggregated interviewer rating of family standard of living at year =* **age** *variable*

*AGGsol\_previous Aggregated interviewer rating of family standard of living from the previous year (i.e at year =* **age\_minus1**)

*1=above average (=1 or 2 from above)*

*2=average (=3 from above)*

*3=below average poor (=4 or 5 from above)*

avsol5 Averaged interviewer rating of family living standards 0-5 years.

As an alternative to the above measures you may prefer to use this summary measure which is the average of interviewer ratings over the 5 years, with this average taken over all non-missing observations. The final scale has been multiplied by 10. Thus the score ranges from 10-50 with a score of 10 implying the family has been consistently rated as affluent and a score of 50 implying the family has been consistently rated as very poor over the 5 year period.

**Measures of Family/Household Structure**

single0 Child entered single parent family at birth

single1 Single parent family/number of parents at 1 yr

single2 Single parent family/number of parents at 2 yr

single3 Single parent family/number of parents at 3 yr

single4 Single parent family/number of parents at 4 yr

single5 Single parent family/number of parents at 5 yr

…

single13 Single parent family/number of parents at 13 yr

All measures *above* are dichotomous, coded as:

1 = Single parent family

2 = Two parent family

*z1single0 0=two parent family 1=single parent family at birth*

*z1single1 0=two parent family 1=single parent family at 1 yr*

*z1single2 0=two parent family 1=single parent family at 2 yr*

*z1single3 0=two parent family 1=single parent family at 3 yr*

*z1single4 0=two parent family 1=single parent family at 4 yr*

*z1single5 0=two parent family 1=single parent family at 5 yr*

*…*

*z1single13 0=two parent family 1=single parent family at 13 yr*

*z1single 0=two parent family 1=single parent family at year=* **age** *variable*

*z1single\_previous 0=two parent family 1=single parent family from the previous year (i.e at year=* **age\_minus1** *variable)*

householdsize1 Total number of people living in household 1 yr

householdsize2 Total number of people living in household 2 yr

householdsize3 Total number of people living in household 3 yr

householdsize4 Total number of people living in household 4 yr

householdsize5 Total number of people living in household 5 yr

…

householdsize13 Total number of people living in household 13 yr

Includes all immediate family (parents, child, child’s siblings) plus any other relatives (eg grandparents, etc) or non-relatives (eg boarders) living in the household. Range 2-14.

*householdsize Total number of people living in household at year =* ***age*** *variable*

*householdsize\_previous Total number of people living in household in previous year (i.e in year =* ***age\_minus1****)*

*z1householdsize1 0=less than 4 people; 1=four or greater number of people 1 yr*

*z1householdsize2 0=less than 4 people; 1=four or greater number of people 2 yr*

*z1householdsize3 0=less than 4 people; 1=four or greater number of people 3 yr*

*z1householdsize4 0=less than 4 people; 1=four or greater number of people 4 yr*

*z1householdsize5 0=less than 4 people; 1=four or greater number of people 5 yr*

*…*

*z1householdsize13 0=less than 4 people; 1=four or greater number of people 13 yr*

*z1householdsize 0=less than 4 people; 1=four or greater number of people at year =* **age** *variable*

*z1householdsize\_previous 0=less than 4 people; 1=four or greater number of people; in previous year (i.e in year =* ***age\_minus1****)*

kids1 Number of children in family 1 yr

kids2 Number of children in family 2 yr

kids3 Number of children in family 3 yr

kids4 Number of children in family 4 yr

kids5 Number of children in family 5 yr

…

kids13 Number of children in family 13 yr

Includes the survey child and their siblings (including all full, half, step, adoptive or foster sibs). Range 1-10.

*kids Number of children in family – during year=***age** *variable*

*kids\_previous Number of children in family– during previous year (i.e when year =***age\_minus1** *variable)*

*z1kids1 0=two or less children; 1= over two children –yr1*

*z1kids2 0=two or less children; 1= over two children –yr2*

*z1kids3 0=two or less children; 1= over two children –yr3*

*z1kids4 0=two or less children; 1= over two children –yr4*

*z1kids5 0=two or less children; 1= over two children –yr5*

*…*

*z1kids13 0=two or less children; 1= over two children –yr13*

*z1kids 0=two or less children; 1= over two children – during year=***age** *variable*

*z1kids\_previous 0=two or less children; 1= over two children – during previous year (i.e when year =***age\_minus1** *variable)*

accom1 Type of accommodation 1 yr

accom2 Type of accommodation 2 yr

accom3 Type of accommodation 3 yr

accom4 Type of accommodation 4 yr

accom5 Type of accommodation 5 yr

…

accom13 Type of accommodation 13 yr

accom Type of accommodation during year=**age** variable

The type of accommodation the family was living in at each year, coded as:

1 = Detached house

2 = Townhouse or ownership flat

3 = Flat (not ownership)

4 = Boarding house

9 = Other (e.g. caravan)

*z1accom1 0=detached house; 1=not a detached house yr 1*

*z1accom2 0=detached house; 1=not a detached house yr 2*

*z1accom3 0=detached house; 1=not a detached house yr 3*

*z1accom4 0=detached house; 1=not a detached house yr 4*

*z1accom5 0=detached house; 1=not a detached house yr 5*

*…*

*z1accom13 0=detached house; 1=not a detached house yr 13*

*z1accom 0=detached house; 1=not a detached house* - during year=**age** variable

*z1accom\_previous 0=detached house; 1=not a detached house - during previous year (i.e when year =***age\_minus1** *variable)*

homeown1 Home owned or rented 1 yr

homeown2 Home owned or rented 2 yr

homeown3 Home owned or rented 3 yr

homeown4 Home owned or rented 4 yr

homeown5 Home owned or rented 5 yr

…

homeown13 Home owned or rented 13 yr

A measure of home ownership at each year, coded as:

1 = Home owned or mortgaged

2 = Rented (private owner)

3 = Rented (state or local authority)

9 = Other (e.g. boarding)

*z1homeown1 0=homeowned/mortaged; 1= not homeowned/mortaged yr 1*

*z1homeown2 0=homeowned/mortaged; 1= not homeowned/mortaged yr 2*

*z1homeown3 0=homeowned/mortaged; 1= not homeowned/mortaged yr 3*

*z1homeown4 0=homeowned/mortaged; 1= not homeowned/mortaged yr 4*

*z1homeown5 0=homeowned/mortaged; 1= not homeowned/mortaged yr 5*

*…*

*z1homeown13 0=homeowned/mortaged; 1= not homeowned/mortaged yr 13*

*z1homeown 0=homeowned/mortaged; 1= not homeowned/mortaged - during year=***age** *variable*

*z1homeown\_previous 0=homeowned/mortaged; 1= not homeowned/mortaged house - during previous year (i.e when year =***age\_minus1** *variable)*

bedrooms1 Number of bedrooms in home 1 yr

bedrooms2 Number of bedrooms in home 2 yr

bedrooms3 Number of bedrooms in home 3 yr

bedrooms4 Number of bedrooms in home 4 yr

bedrooms5 Number of bedrooms in home 5 yr

…

bedrooms13 Number of bedrooms in home 13 yr

The number of bedrooms in the family’s accommodation at each year

*bedrooms Number of bedrooms in home- during year=***age** *variable*

*bedrooms\_previous Number of bedrooms in home- during previous year (i.e when year =***age\_minus1** *variable)*

*z1bedrooms1 0=three or more bedrooms; 1=less than three bedrooms yr 1*

*z1bedrooms2 0=three or more bedrooms; 1=less than three bedrooms yr 2*

*z1bedrooms3 0=three or more bedrooms; 1=less than three bedrooms yr 3*

*z1bedrooms4 0=three or more bedrooms; 1=less than three bedrooms yr 4*

*z1bedrooms5 0=three or more bedrooms; 1=less than three bedrooms yr 5*

*z1bedrooms13 0=three or more bedrooms; 1=less than three bedrooms yr 13*

*z1bedrooms 0=three or more bedrooms; 1=less than three bedrooms - during year=***age** *variable*

*z1bedrooms\_previous 0=three or more bedrooms; 1=less than three bedrooms house - during previous year (i.e when year =***age\_minus1** *variable)*

*overcrowd* householdsize/bedrooms;

*z1overcrowd 0=overcrowd<2; 1=overcrowd>=2*

*z1overcrowd\_previous 0=overcrowd<2; 1=overcrowd>=2 - during previous year (i.e when year =***age\_minus1** *variable)*

**Measures of Family Functioning**

chpar1 Changes of parents 0-1 yr

chpar2 Changes of parents 1-2 yr

chpar3 Changes of parents 2-3 yr

chpar4 Changes of parents 3-4 yr

chpar5 Changes of parents 4-5 yr

…

chpar13 Changes of parents 12-13 yr

Dichotomous indicators of whether the child experienced a change of parents in a given 12 month period from birth to age 13 years. A change of parents could include parental separation, reconciliation, remarriage, entry of a step parent, placement with foster parents, parental death, or any other change of custodial parents. Scored as:

1 = Child experienced a change of parents in year

2 = Child did not experience a change of parents

*z1chpar1 0=did not have change in parents;1=did have change in parents -yr 1*

*z1chpar2 0=did not have change in parents;1=did have change in parents -yr 2*

*z1chpar3 0=did not have change in parents;1=did have change in parents -yr 3*

*z1chpar4 0=did not have change in parents;1=did have change in parents -yr 4*

*z1chpar5 0=did not have change in parents;1=did have change in parents -yr 5*

*…*

*z1chpar13 0=did not have change in parents;1=did have change in parents -yr 13*

*z1chpar 0=did not have change in parents;*

*1=did have change in parents -during year=****age*** *variable*

*z1chpar\_previous 0=did not have change in parents;*

*1=did have change in parents –during previous year (i.e when year was* **age\_minus1***)*

chres1 Changes of residence 0-1 yr

chres2 Changes of residence 1-2 yr

chres3 Changes of residence 2-3 yr

chres4 Changes of residence 3-4 yr

chres5 Changes of residence 4-5 yr

…

chres13 Changes of residence 12-13 yr

Number of changes of residence experienced by the child in each 12 month period from birth to age 5 years.

*chres Changes of residence , during year =* **age** *variable*

*chres\_previous Changes of residence , during previous year (i.e during* ***age\_minus1*** *year*)

*Z1chres1 0=no change in residence; 1=one or more changes in residence - yr 1*

*Z1chres2 0=no change in residence; 1=one or more changes in residence - yr 2*

*Z1chres3 0=no change in residence; 1=one or more changes in residence - yr 3*

*Z1chres4 0=no change in residence; 1=one or more changes in residence - yr 4*

*Z1chres5 0=no change in residence; 1=one or more changes in residence - yr 5*

*…*

*Z1chres13 0=no change in residence; 1=one or more changes in residence - yr 13*

*Z1chres 0=no change in residence; 1=one or more changes in residence during year =* **age** *variable*

*Z1chres\_previous 0=no change in residence; 1=one or more changes in residence - during previous year (i.e during* ***age\_minus1*** *year*)

stress2 Number of adverse family life events 1-2 yr

stress3 Number of adverse family life events 2-3 yr

stress4 Number of adverse family life events 3-4 yr

stress5 Number of adverse family life events 4-5 yr

We did not get stress for years 6 to 13

*stress* *Number of adverse family life events* *living at year =* **age** *variable*

*stress\_previous* *Number of adverse family life events* *from the previous year (i.e at year =* **age\_minus1**)

A measure of exposure to adverse family life events in each year from age 1-2 to age 4-5, based on the Holmes and Rahe Social Readjustment Scale. Note this measure is not available for the period 0-1 year. (However, family life events are likely to be amongst the strongest predictors of medical contacts for accidents)

**Burt Word Reading Test (8-13 years)**

At each year from age 8-13 children’s reading ability was assessed using the Burt Word Reading Test. **Note**: **Testing was limited to children resident in the Canterbury region at each age. This sample represented approximately 80% of the total sample available for assessment at each year.**

burt8 Burt Word Reading Test score 8 yrs

burt9 Burt Word Reading Test score 9 yrs

burt10 Burt Word Reading Test score 10 yrs

burt11 Burt Word Reading Test score 11 yrs

burt12 Burt Word Reading Test score 12 yrs

burt13 Burt Word Reading Test score 13 yrs

burtImp burt with values imputed for those with missing values. This is the variable actually used for analysis

**OTHER VARIABLES**

**Early childhood education**

NPRESCH The number of years of preschool education received by the child (Range 0-3)

**Measures of early parenting style**

The following two measures were derived from the Home Observation for Measurement of the Environment (HOME) inventory (Elardo, Bradley and Caldwell 1983). Emotional responsiveness measures the extent to which the mother is attuned to the child’s needs and responds in a warm and loving manner. A high score implies high emotional responsiveness. Maternal punitiveness measures the extent to which the mother utilized a restrictive and punitive parenting style. A high score implies greater punitiveness. Both measures were assessed on the basis of interviewer observations averaged over the period from age 3-5 yrs

INTERACT Maternal emotional responsiveness (Range 0-10)

PUNISH Maternal punitiveness (Range 0-5)

**Childhood conduct problems**

At each year from age 6-10 years maternal and teacher reports of the child’s tendencies to conduct disordered and oppositional behaviours were obtained. At age 6 years these reports were based on the Rutter parent and teacher report scales. At ages 7-10 the reports were based on scales that combined the Rutter and Conners parent or teacher report scales. (See Fergusson, Horwood & Lloyd, *Journal of Child Psychology and Psychiatry* 1991; 32: 257-274 for detailed description of scale items). Scale items were scored on a 3-pt scale (doesn’t apply, applies somewhat, definitely applies) and the total scale scores are based upon a simple item summation for each reporting source at each age.

mcond6 Maternal report of conduct problems 6 years (Range 9-25)

mcond7 Maternal report of conduct problems 7 years (Range 21-55)

mcond8 Maternal report of conduct problems 8 years (Range 21-57)

mcond9 Maternal report of conduct problems 9 years (Range 21-53)

mcond10 Maternal report of conduct problems 10 years (Range 21-59)

tcond6 Teacher report of conduct problems 6 years (Range 9-27)

tcond7 Teacher report of conduct problems 7 years (Range 20-54)

tcond8 Teacher report of conduct problems 8 years (Range 20-54)

tcond9 Teacher report of conduct problems 9 years (Range 20-57)

tcond10 Teacher report of conduct problems 10 years (Range 20-59)

cond The average of mcond and tcond

**Derived Variables Used in the Simulation**

**Sptype** – in dataset this is derived from the change from z1single\_previous to z1single:

1='became partnered'

2='broke up'

3='same status as b4';

**Typeofchange** – in dataset this is derived from the change from previous parental ages variables (whether 99 or not) to current parental ages variables (whether 99 or not). For the purpose of creating and using typeofchange variable in the simulation, I put the -2 category in as a number 5, as -2 is really a rare subset of number 5.

-**2**='diffkindparent in singlep fam'

**0**='both father and mother left'

**1**='got new mother/ mother back'

**2**='got new father/ father back'

**3**='mother left'

**4**='father left'

**5**='have same number of parents'

**Mumgroup**

**0=** no mum

**1=** birthmother

**2=** not birthmother not same mum as last year

**3=** not birthmother but same mum as last year

**Dadgroup**

**0=** no dad

**1=** birthfather

**2=** not birthfather not same dad as last year

**3=** not birthfather but same dad as last year

**fage\_imputed**

Already provided in the basefile, this is a continuous version of **fage**, the father’s age at the birth of the child. It is derived via the program ‘imputing birth fathers age for oliver - sep011.sas’ which uses, depending on the circumstances, a combination of father’s age at year 1; mother age at birth of child; and average ages of fathers in the unbounded groups, to impute father’s continuous age at birth of child.