

1 [Supplementary materials](#)

2 **Covariate dictionary.** Detailed descriptions of adversity exposure measures and data sources

3 **Supplementary Table S1.** Data coverage for constructed adversity exposures. Development sample.

4 **Supplementary Table S2.** Adversity exposures for children in development sample excluded or
5 censored due to death. Counts and percentages.

6 **Supplementary Table S3.** Exposure to adversity in validation samples (Number and percentage).

7 **Supplementary Table S4.** Alternative model parameterisations for risk of developing schizophrenia
8 following exposure to adversity across domains A1 – Discontinuity in parenting and A4 – Area level
9 socio-economic/demographic environment. Hazard ratios and 95% Confidence intervals.

10 **Supplementary Table S5.** Observations excluded from final model parameter estimation due to large
11 influence. N=27 observations, all with schizophrenia diagnosis. Adversity exposure comparison with
12 non-excluded observations.

13 **Supplementary Table S6.** Parameter estimates derived from development sample including
14 observations with large influence. Comparison with optimum models

15 **Supplementary Table S7.** Elementary validations for models. Robustness of estimated Cox model
16 hazard ratios - comparison with estimated logistic regression odds ratios (Ratios and 95% Confidence
17 intervals). Proportional hazards validity – test for non-zero correlation of Schoenfeld residuals with
18 time (p values).

19 **Supplementary Table S8.** Risk of developing schizophrenia following exposure to adversity across
20 five domains. Alternative Global models. Hazard ratios and 95% Confidence intervals.

21 **Supplementary Table S9.** Alternative Global Cox model parameterisations describing risk of
22 schizophrenia following exposure to adversity. Harrell's Concordance statistics.

23 **Supplementary Figure S1.** Distributions of predicted log hazard ratios for five domain models and
24 Global model, derived from development sample.

25

1 Covariate dictionary

2 1. Covariates in manuscript

3 (See also section 3. which describes the larger pool of candidate variables from which the final
4 variables were chosen).

5 A1 - Discontinuity in parenting before the age of 10 years covered any separation of a child from a
6 parent categorised by the child's developmental epoch: i) under one year of age; ii) ages one to four;
7 and iii) ages five to nine.

- 8 • *Hospitalisations of the child, the mother or the father* (sourced from the [Hospital Morbidity
9 Data Collection](#)) were classified as adverse if multiple admissions which totalled more than 8
10 days occurred within in any developmental epoch. To exclude a child's birth related hospital
11 stay being included as separations between mother and child, the less than 1 year period
12 was defined to begin at age 15 days for the child and their mother, rather than 0 days as for
13 the fathers.
- 14 • *Parental deaths* before a child's 10th birthday were sourced from the [Western Australian
15 Death Registry](#).
- 16 • The [Department of Communities, Child Protection and Family Support Division database](#)
17 provided information pertaining to any child *placement in foster care* before their 10th
18 birthday.
- 19 • Custodial sentences leading to *incarceration* of a child's mother or father during each of a
20 child's developmental epochs were determined from the [Western Australian Department of
21 Justice](#).

22

23 A2 - Family functioning

- 24 • *Corrective services contacts* for a child's mother or father, before the child's 10th birthday,
25 were obtained from the [Western Australian Department of Justice](#). As well as incarceration,
26 these contacts include diversionary programmes and non-custodial orders.
- 27 • *Child protection contacts* before a child's 10th birthday covering any notification of
28 allegations of child abuse (sexual, physical and emotional) and neglect were sourced from
29 the [Department of Communities, Child Protection and Family Support Division database](#)
30 and were classified as adverse at two levels i) if, before a child's 10th birthday, any
31 notification of abuse was made related to a child or any of their maternal siblings but with
32 none of those notifications substantiated and ii) as for i), but with at least one notification of
33 abuse being substantiated.

34 A3 - Family structure

- 35 • *mother's age* at time of child's birth was calculated using the [Midwives' Notification
36 System](#), and categorised as less than 20 years or at least 20 years.
- 37 • *paternal age* at child's birth was calculated for those fathers with valid [Birth Registration
38 details](#) and categorised as less than 25 years, at least 25 years but less than 45 years, at least
39 45 years.

- 1 • *marital status* (partnered / unpartnered) of the child's mother at the time of a child's birth
- 2 and the
- 3 • *child's birth order* (1st, 2nd or 3rd compared to 4th or greater) amongst maternal siblings
- 4 were determined from the [Midwives' Notification System](#).
- 5 • *size of the family* (2 children, 5 or more children, or 1, 3 or 4 children) a child belonged to at
- 6 the time of their 10th birthday was calculated by combining birth order with any [Hospital](#)
- 7 [Morbidity Data Collection](#) entries after 2001 which identified further live births to a child's
- 8 mother.

9 A4 - Area level socio-economic/demographic environment

10 The Australia Bureau of Statistics uses a hierarchical geographic classification system {Australian
 11 Bureau of Statistics, 2006 #49} for collection and publication of official statistics. The finest level at
 12 which statistics were published for the 2001 Census was the Census Collection District (CCD), a
 13 grouping of approximately 200 geographically proximate households. The CCD corresponding to a
 14 mother's address at the time of the child's birth was provided by the Data Linkage Branch which was
 15 matched to the relevant area level 2001 census statistics. In a minority of cases where CCD level
 16 statistics were not available, aggregation at postcode level CCD was used.

- 17 • *Socioeconomic status* was determined using the [Australian Bureau of Statistics Socio-](#)
- 18 [Economic Indexes for Areas](#) {Australian Bureau of Statistics, 2008 #48}. We used the Index
- 19 of Relative Socioeconomic Disadvantage, comparing those children born in areas of the
- 20 quintile of greatest disadvantage to all other children. The index used was that for the
- 21 available census (1986, 1991, 1996 or 2001) closest in time to the date of birth.
- 22 Transformation of index values to quintiles allowed comparison across subject birth years.
- 23 • *Remoteness or level of urbanicity* of the geographical area of a mother's residence at the
- 24 time of her child's birth was determined using the [Australian Standard Geographic](#)
- 25 [Classification—Remoteness Area](#), as defined in 2001. {Australian Bureau of Statistics, 2006
- 26 #49}. We compared those children born to mothers resident in Remote or Very Remote
- 27 areas to children born to mothers resident in all other areas.

28 Further area level measures determined from 2001 census information were percentage of persons
 29 in a CCD who:

- 30 • identified as *Indigenous* (of Aboriginal or Torres Strait Islander descent) (<25% or >=25%)
- 31 • were *Australian born* (>=75% or <75%)
- 32 • *never married* (<50% or >=50%)
- 33 • lived in *one parent families* (<20% or >=20%)
- 34 • were *unemployed* (<5% or >=5%)
- 35 • lived in different *residence to 5 years* previous (<60% or >=60%)
- 36 • lived in different *residence to 1 year* previous (<30% or >=30%)
- 37 • lived in *semi-detached* dwellings (<10% or >=10%)
- 38 • lived in *flats* (<5% or >=5%)
- 39 • lived in *rented* dwellings (<30% or >=30%)
- 40 • had *no post school qualifications* (<40% or >=40%)
- 41 • *spoke a language other than English at home and did not speak English well* (>=2% or <2%).
- 42

- 1 • *Ethnic heterogeneity* was calculated for each CCD using the formula: $E = 1 - P_1^2 - P_2^2$ where P_1
2 is the percentage of residents in that CCD who identified as Indigenous, and P_2 is the
3 percentage of residents in that CCD who spoke a language other than English at home and
4 who did not speak English well. {Morgan, 2008 #149} This measure is at its lowest when the
5 diversity within a population is highest. Areas of high heterogeneity were defined as those
6 where E was calculated to be less than 0.99.
- 7 • *Total area-level crime* was based on the arrest rate per 1000 residents in a Local
8 Government Area during 2002, as provided by the [Western Australian Crime Research](#)
9 [Centre](#). Children born in areas where at least 50 arrests per 1000 residents were recorded
10 were compared to all other children.
- 11 • *socioeconomic inequality* for a postcode was derived from the [standard deviation](#) of the
12 distribution of the [Index of Relative Socioeconomic Disadvantage](#) scores for CCDs within
13 that postcode. Across all postcodes, the distribution of standard deviations was categorised
14 with the quartile of smallest standard deviations defined as minimal localised inequality.

15 A5 - Family level sociodemographic status

- 16 • A binary variable was created to indicate whether the child's *known father* had been named
17 on the State [Birth Registration](#) record.
- 18 • The *child's Indigenous* status was scored positive if the child and/or either parent was
19 identified as Indigenous (of Aboriginal or Torres Strait Islander descent) in [any of the data](#)
20 [sources available](#).
- 21 • Maternal and paternal place of birth were extracted from State [Birth Registration](#) records.
22 In turn, *Migrant status* of parents was classified according to the affluence of the country of
23 origin, using [World Bank Income Categories](#), and for Australian born parents, whether they
24 had been born within or outside Western Australia. Based on patterns of response in
25 preliminary models, children of Western Australian born parents were grouped with children
26 of parents born in lower income countries
- 27 • The State [Birth Registration's](#) record of a *father's occupation* at the time of a child's birth
28 was classified according to skill level based on the [Australian Bureau of Statistics](#)
29 [Occupation Classifications](#). Categorisation was defined as fathers in employment requiring
30 middle to high skill level compared to all other employment statuses such as low skill,
31 unemployed, not in labour-force etc).

1 2. Descriptions of Data sources

2

3 **Hospital Morbidity Data Collection**

4 Records available 1st January 1980 to 31st December 2017

5 The collection is managed by the WA Health System and records all episodes of care that occur in
6 public and licensed private acute and psychiatric hospitals as well as licensed private day
7 surgeries. Excluded episodes of care include those pertaining to patients attending emergency,
8 outpatient or community health services, those in residential aged care facilities, still births and
9 mothers electing to deliver at home

10

11 **Mental Health Information System**

12 Records available 1st July 1966 to 31st December 2015.

13 Covers records of all public and private inpatient hospital admissions, as well as public outpatient
14 and ambulatory care contacts with mental health services across the State.

15

16 **Western Australian Death Registry**

17 Records available 1st January 1980 to 19th July 2018 for all deaths registered in Western Australia

18 **Department of Communities, Child Protection and Family Support Division database**

19 Records available 1st July 1989 to 29th February 2016 for notifications.

20 Records available 1st April 1981 to 12th May 2016 for placements in care.

21

22 The Department of Communities, Child Protection and Family Support Division record all
23 notifications of child abuse (sexual, physical and emotional) and neglect meeting standard
24 Australian definitions {Bromfield, 2008 #576;Bromfield, 2008 #576} and the outcome of
25 subsequent investigations – whether substantiated or not. The date of reporting does not always
26 reflect the date of abuse, as reporting of the abuse may happen years after the event/s. The
27 Division also records all placements of children in foster care and the duration of such
28 placements.

29

30 **Western Australian Department of Justice**

31 Records available 1st January 1980 to 15th October 2011.

32 The Western Australian Department of Justice holds electronic records dating back to the early
33 1970s. The data include information on youth detention, adult prison populations and non-
34 custodial orders (adult community-based and youth justice orders). They cover Western

1 Australia's 14 public prisons, two privately run prisons, five work camps and one detention
2 centre, as well as offenders on probation, parole, other community orders and diversionary
3 programmes for young people. The minimum age for inclusion on the Western Australian
4 corrective services records is 10 years of age.

5 **Midwives Notification System**

6 Records available 1st January 1980 to 31st December 2001.

7 The Midwives Notification system {Gee, 1994 #37} includes mandatory, prospectively collected
8 data on all infants born in Western Australia at 20 weeks gestation or more, or weighing at least
9 400 grams, including home births, and not restricted to live births. Demographic information for
10 the mother includes age, date of birth, marital status, postcode of residence and ethnic identity.
11 (Outcomes of previous pregnancies, occurrence of any complications, medical conditions or
12 treatments during pregnancy and details of labour including any complications or interventions
13 are also recorded). Vital details pertaining to the baby include date and time of birth, gender,
14 whether baby was one of a multiple birth event and if baby was stillborn or alive.

15 **Birth Registration records – mother and father details**

16 Records available for births 1st January 1980 to 31st December 2001.

17 In Western Australia, a Birth Registration Form must be lodged with the Registry of Births Deaths
18 and Marriages within 60 days of the birth of a child to enable the newborn to be officially
19 recognised for a range of services. This must be completed and signed by both parents. If only
20 one parent has signed, they must provide an explanation why the other parent has not signed.
21 Information requested of each parent includes Indigenous identity, age, date and place of birth,
22 and occupation.

23 **Australian Bureau of Statistics Socio-Economic Indexes for Areas (Index of Relative Socioeconomic 24 Disadvantage)**

25 The Socio-economic Indexes for Areas (SEIFA) is a set of indexes created by the Australian
26 Bureau of Statistics from census data. For general information about all SEIFA, including SEIFA_A
27 = Index of Relative Socio-economic Disadvantage, see
28 <https://www.abs.gov.au/ausstats/abs@.nsf/mf/2039.0> .

29 SEIFA_A is constructed differently for each Census, however this study used quintiles of SEIFA_A
30 to assign comparable values for the SEIFA_A index for each census year.

| 31 | <u>Census</u> | <u>Birth years</u> |
|----|---------------|--------------------|
| 32 | 1986 | 1980 – 1987 |
| 33 | 1991 | 1988 – 1993 |
| 34 | 1996 | 1994 – 1998 |

35 **Australian Standard Geographic Classification—Remoteness Area**

36 Remoteness Area categories are based on the Accessibility/Remoteness Index of Australia (ARIA)
37 maintained by the National Centre for Social Applications of GIS (GISCA) at the University of

1 Adelaide. Its calculation includes consideration of the road distance of a location from the
2 nearest population centres providing access to goods and services, taking into account
3 population size. Since the 2001 census, the Australian Bureau of Statistics (ABS) has defined a
4 'Remoteness Area classification Index (RA)' for each non-overlapping Census Collection District
5 (CD), based on the methodology of the ARIA. Each Collection District is then categorised,
6 according to the value of the Index as one of 5 levels of urbanicity: 'Major Cities of Australia',
7 'Inner Regional Australia', 'Outer Regional Australia', 'Remote Australia', and 'Very Remote
8 Australia'.

9 The ABS RA measure is superior to a postcode-only based definition of urbanicity due to
10 increased geographical precision and rigor of definition. A limitation of using the RA classification
11 correct at Census 2001 is the omission of historical RA trends. For example, the residence of the
12 mother of a child born in 1980 may have been considered Inner Regional at this time. With
13 increased urban spread, the very same residence may be in a Major Cities area in 2001. The
14 allocation of the 2001 RA value to the residence of that child does not acknowledge that for at
15 least part of the child's early life, the residential environment was considered Inner Regional.

16 **Western Australian Crime Research Centre**

17 The Crime Research Centre of the University of Western Australia was established in 1989 as
18 a major and innovative contributor to the collection and analysis of statistical data on crime in
19 Western Australia. It published annual crime and justice statistics which enhanced the
20 understanding of the extent of the crime problem in the state and trends in criminal justice. The
21 arrest rate per 1,000 residents in a Local Government Area during 2002 were provided to the
22 authors for the specific requirements of this study.

23 **World Bank Income Category rankings**

24 The World Bank Income Category list of global economies (July 2012) was accessed 5th March
25 2013. <http://siteresources.worldbank.org/DATASTATISTICS/Resources/CLASS.XLS>. Based on Gross
26 National Income figures for 2011, each economy was classified as Low Income, Lower middle
27 income, Upper middle income, or High income. The Standard Australian Classification of
28 Countries (SACC) - Australian Bureau of Statistics was used to link countries to economies.
29 [http://www.abs.gov.au/ausstats/abs@.nsf/Latestproducts/3E3848A2DCF7FF4FCA25744B001535](http://www.abs.gov.au/ausstats/abs@.nsf/Latestproducts/3E3848A2DCF7FF4FCA25744B001535DB?opendocument)
30 [DB?opendocument](http://www.abs.gov.au/ausstats/abs@.nsf/Latestproducts/3E3848A2DCF7FF4FCA25744B001535DB?opendocument)

31 **Australian Bureau of Statistics occupation classifications**

32 The Australian Bureau of Statistics publication, *1220.0 - Australian Standard Classification of*
33 *Occupations (ASCO) Second Edition, 1997* classifies 986 unique occupations into 9 Major Groups
34 which map to 5 ordinal Skill Levels.

35 [http://www.ausstats.abs.gov.au/ausstats/free.nsf/0/A86A0162E6F672DFCA256ADB001D10D4/\\$](http://www.ausstats.abs.gov.au/ausstats/free.nsf/0/A86A0162E6F672DFCA256ADB001D10D4/$File/asco.pdf)
36 [File/asco.pdf](http://www.ausstats.abs.gov.au/ausstats/free.nsf/0/A86A0162E6F672DFCA256ADB001D10D4/$File/asco.pdf)

37

3. Raw constructs considered in preliminary screening for association with schizophrenia.

Before arriving at the covariates referenced in section 1, we considered a wider range of adversity exposures that were considered to have potential to measurably impact the risk of schizophrenia diagnosis. Where the register data provided sufficient scope and accuracy for their calculation, we used a pragmatic, data-driven approach, to construct categorical variables. Categories were designed to retain maximum information about the relationship between the candidate variable and schizophrenia risk, while summarising the relationship parsimoniously and providing for the possible creation of equivalent variables by other researchers. Thus, while we acknowledge that information can be lost by categorising a continuous measure, we anticipated that linear relationships would be insufficient to account for likely complex associations between measures of adversity and schizophrenia. Relationships were initially assessed with several categories so that general functional forms could be identified from the data (rather than assumed prior) and then piecewise functions could be constructed to approximate the functional forms. In summary, we assessed bivariate associations between categorical variables and schizophrenia (specifically Nelson-Aalen graphs and hazard ratios) to i) exclude those variables with insufficient evidence of association from further consideration and ii) for those variables with evidence of association, to identify where categories could be collapsed without substantial loss of explanation of varying risk. The full range of variables considered, from which final manuscript variables were derived, are listed below.

A1 - Discontinuity in parenting before the age of 10 years covered any separation of a child from a parent categorised by the child's developmental epoch: (i) under one year of age; (ii) ages one to four; and (iii) ages five to nine.

- Numbers of, and length of *Hospitalisations* of the child, the mother or the father (sourced from the Hospital Morbidity Data Collection) were classified as: no recorded hospitalisations, single day procedure/s only, single admission of any length of at least overnight, multiple admissions totalling 1-7 nights, multiple admissions totalling at least 8 nights. To exclude a child's birth related hospital stay being included as separations between mother and child, the less than 1 year period was defined to begin at age 15 days for the child and their mother, rather than 0 days as for the fathers.
- *Parental deaths* were sourced from the Western Australian Death Registry separately for mothers and fathers and, depending on date of death, attributed to the corresponding epoch of the child's development.
- The Department of Communities, Child Protection and Family Support Division database was searched for any child *placement in foster care*. Numbers of and length of placements were recorded for each of a child's developmental epochs.
- Custodial sentences leading to *incarceration* of a child's mother or father (number of and durations) were determined from the Department of Justice records. Tallies were considered separately for each epoch of the child's development.

A2 - Family functioning

The Department of Communities, Child Protection and Family Support Division record notifications of allegations of child abuse (sexual, physical and emotional) and neglect meeting standard Australian definitions {Bromfield, 2008 #576; Bromfield, 2008 #576} and the outcome of subsequent

1 investigations – whether substantiated or not. It should be noted that the date of notification can
2 occur well after the date the alleged abuse took place, and, due to the nature of the trauma
3 surrounding the allegations, the accuracy of the dates of alleged offences is hard to establish. Hence,
4 our inclusion of only notifications received up to a child's 10th birthday will necessarily exclude some
5 instances of abuse which had occurred but had not yet been reported. Recognising the possibility of
6 trauma by association, we considered any occurrences of child protection contact relating to a
7 child's maternal siblings as well as those relating specifically to the child.

- 8 • The number of *Child protection contacts* before a child's 10th birthday were classified as: no
9 allegations made relating to child or any of their maternal siblings, number of allegations
10 relating to any sibling, number of substantiated allegations related to any sibling, number of
11 allegations related to child, number of substantiated allegations related to child.

12 Criminal offending data were obtained from the Western Australian Department of Justice and, as
13 well as detention, include information on diversionary programmes and non-custodial orders.

- 14 • The number of any such corrective services contacts for a child's mother or father, before
15 the child's 10th birthday, were noted.

16

17 A3 - Family Structure

18 Using the Midwives' Notification System,

- 19 • *mother's age* at time of child's birth was calculated and categorised as less than 20 years,
20 20-24 years, 25-29 years, 30-34 years or at least 35 years.
- 21 • Maternal *marital status* (never married, widowed, divorced, separated, married (including
22 defacto), other status) at the time of a child's birth was extracted,
- 23 • both the child's *birth order* amongst maternal siblings and
24 • whether they were part of a *multiple birth* event
25 were determined. This data, combined with any Hospital Morbidity Data Collection entries after
26 2001 which identified further live births to a child's mother
- 27 • allowed calculation of the *size of the family* a child belonged to at the time of their 10th
28 birthday.
- 29 • For those fathers with valid birth registration details, *paternal age* at child's birth was
30 calculated and categorised as for mothers with further refinement for 35-39 years, 40-44
31 years and at least 45 years.

32 A4 - Area level socio-economic/demographic environment

33 The Australia Bureau of Statistics uses a hierarchical geographic classification system {Australian
34 Bureau of Statistics, 2006 #49} for collection and publication of official statistics. The finest level at
35 which statistics were published for the 2001 Census was the Census Collection District (CCD), a
36 grouping of approximately 200 geographically proximate households. The CCD corresponding to a
37 mother's address at the time of the child's birth was provided by the Data Linkage Branch which
38 allowed access to relevant area level 2001 census statistics. In a minority of cases where CCD level
39 statistics were not available, aggregation at postcode level CCD was used.

- 1 • *Socioeconomic status* was determined using the Australian Bureau of Statistics Socio-
2 Economic Indexes For Areas {Australian Bureau of Statistics, 2008 #48}. We used the Index
3 of Relative Socioeconomic Disadvantage. The index used was that for the available census
4 (1986, 1991, 1996 or 2001) closest in time to the date of birth. Transformation of index
5 values to quintiles allowed comparison across subject birth years.
- 6 • The geographical *remoteness or level of urbanicity* of the mother's residence at time of a
7 child's birth was determined using the Australian Standard Geographic Classification—
8 Remoteness Area, as defined in 2001. {Australian Bureau of Statistics, 2006 #49}.
9 Remoteness Area categories are based on the road distance of a location from the nearest
10 population centres providing access to goods and services, taking into account population
11 size. They are classified as: Major City, Inner Regional, Outer Regional, Remote and Very
12 Remote.

13 It is noted that both Socioeconomic status and Remoteness are composite measures, which may
14 intrinsically incorporate some information measured by the following additional area level
15 measures determined from 2001 Australian Bureau of Statistics Census Data. As such, additional
16 census data variables were assessed for association with schizophrenia after adjustment for
17 both Socioeconomic status and Remoteness. Alternative distributional categorisations were
18 considered: i) approximately 3 categories corresponding to natural cut points of the relevant
19 distribution, and ii) quintiles for each of the variables 'percentage of persons in a CCD who: '

- 20 • identified as *Indigenous* (of Aboriginal or Torres Strait Islander descent), were *Australian*
21 *born, never married, lived in one parent families, were unemployed, lived in the same*
22 *residence as 5 years previous, lived in the same residence as 1 year previous, lived in semi-*
23 *detached dwellings, lived in flats dwellings, lived in rented dwellings, had no post school*
24 *qualifications, or spoke a language other than English at home and did not speak English*
25 *well.*
- 26 • *Ethnic heterogeneity* was calculated for each CCD using the formula: $E = 1 - P_1^2 - P_2^2$ where
27 P_1 is the percentage of residents in that CCD who identified as Aboriginal or Torres Strait
28 Islander, and P_2 is the percentage of residents in that CCD who spoke a language other than
29 English at home and who did not speak English well. {Morgan, 2008 #149} This measure is at
30 its lowest when the diversity within a population is highest. Three heterogeneity categories
31 were defined according to natural cut points in the distribution of E, i) lowest diversity:
32 $0.999 \leq E \leq 1$, ii) some diversity: $0.99 \leq E < 0.999$, and iii) greatest diversity: $0 \leq E < 0.99$.
- 33 • *Area-level crime total, against person and other* was based on the arrest rate per 1000
34 residents in a Local Government Area during 2002 and considered as both quintile and
35 natural cut point categorisations of the distributions.
- 36 • For each postcode, the distribution of the Index of Relative Socioeconomic Disadvantage -
37 Census 2001 scores (one score for each CCD within the postcode) was summarised by its
38 standard distribution. Across all postcodes, the distribution of standard deviations was
39 categorised into quartiles to provide a measure of localised *socioeconomic inequality*.
40

41 42 A5 - Family level sociodemographic status

- 1 • A binary variable was created to indicate whether the child's *known father* had been named
- 2 on the State Birth Registration record.
- 3 • The *child's Indigenous* status was scored positive if the child and/or either parent was
- 4 identified as Indigenous (of Aboriginal or Torres Strait Islander descent) in any of the data
- 5 sources available.
- 6 • Maternal and paternal place of birth were extracted from State Birth Registration records.
- 7 *Migrant status* of parents was classified according to the affluence of the country of origin,
- 8 using World Bank Income categories , and for Australian born parents, whether they had
- 9 been born within or outside Western Australia. Place of birth was then categorised as born
- 10 in Western Australia, born elsewhere in australia, born overseas in a country of low income
- 11 or born overseas in a country of high income
- 12 • The State Birth Registration's record of a *father's occupation* was classified according to skill
- 13 level based on the Australian Bureau of Statistics (ABS) occupation classifications. The ABS
- 14 uses a 5 skill level categorisation which we collapsed to 3 levels: Levels 1 and 2 into
- 15 'Generally higher skilled', Skill Levels 3 and 4 to 'Mid-range skills', and Skill Level 5 as
- 16 'Generally lower skilled', while retaining separate categories for 'Not in Labour force' and
- 17 'Unable to be determined'. Father's occupation status was missing for 3.7% of known
- 18 fathers (6,186/165,754).

19

20

21 **References**

22 **Australian Bureau of Statistics** (2006). Statistical Geography Volume 1 - Australian Standard
 23 Geographical Classification (ASGC), Jul 2006. Catalogue no. 1216.0. Australian Bureau of
 24 Statistics: Canberra.

25 **Australian Bureau of Statistics** (2008). Socio-Economic Indexes for Areas (SEIFA) - Technical
 26 Paper 2006. Catalogue no. 2039.0.55.001. Australian Bureau of Statistics: Canberra.

27 **Bromfield, L. & Holzer, P.** (2008). A National Approach for Child Protection Project Report. .
 28 Australian Institute of Family Studies. : Melbourne.

29 **Gee, V. & Dawes, V.** (1994). Validation Study of the Western Australian Midwives'
 30 Notification System 1992. Occasional Paper 59. p. 18. Health Department of Western
 31 Australia.

32 **Morgan, V. A., Morgan, F., Clare, J., Valuri, G., Woodman, R., Ferrante, A., Castle, D. &**
 33 **Jablensky, A.** (2008). *Schizophrenia and offending: area of residence and the impact of social*
 34 *disorganisation and urbanicity*. Australian Institute of Criminology.

35

Supplementary Table S1 – Data coverage for constructed adversity exposures. Development sample.

| Adversity exposure | Total development sample N=171,588 |
|--|--|
| Discontinuity in Parenting | |
| Mother hospitalised >= 8 days when child aged 15 days <1 year | Complete; some censoring if mother (n=42) deceased <1 yr |
| Mother hospitalised >= 8 days when child aged 1 <5 years | Complete; some censoring if mother (n=42+208) deceased <5 yr |
| Mother hospitalised >= 8 days when child aged 5 years <10 years | Complete; some censoring if mother (n=42+208+389) deceased <10 yr |
| Child hospitalised >= 8 days when child aged 15 days <1 year | Complete; some censoring if child deceased < 1 yr |
| Child hospitalised >= 8 days when child aged 1 <5 years | Complete; some censoring if child deceased < 5 yr |
| Child hospitalised >= 8 days when child aged 5 years <10 years | Complete; some censoring if child deceased < 10 yr |
| Father hospitalised >= 8 days when child aged 0 days <1 year | Complete; some censoring if father (n=129) deceased < 1 yr |
| Father hospitalised >= 8 days when child aged 1 <5 years | Complete; some censoring if father (n=129+499) deceased < 5 yr |
| Father hospitalised >= 8 days when child aged 5 years <10 years | Complete; some censoring if father (n=129+499+813) deceased < 10 yr |
| Deceased parents, either or both, before child aged <10 years | Complete |
| Any placement in foster care before child aged <10 years | Records available FROM 1 st April 1981. Affected N=8,343 ¹ |
| Any incarcerations for mother when child aged <10 years | Records available UNTIL 15 th October 2011. Censored N=1,572 some censoring if mother deceased < 10 yr |
| Any incarcerations for father when child aged 0 days <1 year | Complete; some censoring if father deceased < 1 yr |
| Any incarcerations for father when child aged 1 <5 years | Complete; some censoring if father deceased < 5 yr |
| Any incarcerations for father when child aged 5 years <10 years | Records available UNTIL 15 th October 2011. Censored N=1,572 or if father deceased < 10 yr |
| Family Functioning | |
| Child protection contact before child aged 10 years: | |
| Child or sibling victim of >=1 substantiated abuse | Records available FROM 1 st July 1989. Affected N=70,474 ² |
| Child or sibling subject of >=1 unsubstantiated notification, no substantiated cases | |
| Neither child nor siblings subject of any notifications | |
| Any parental corrective services contact before child aged 10 years | Records available UNTIL 15 th October 2011. Censored N=1,572 |

¹ Children born before 1st April 1981 have censored observations for very early life. Any foster placements after this date, up to a child's 10th birthday, will still contribute to a recorded exposure for that child.

² No cases have completely missing data. Eg, for a child born in 1988, the only missing data is for their 1st year of life. Any substantiations after that, up to their 10th birthday, still contribute to a recorded exposure for that child. Additionally, maternal sibling substantiations contribute to the index child's exposure. It is estimated we are missing ~19.8% of total expected flags, based on numbers observed for children born 1989 onwards.

Or if mother or father deceased <10 yr

Family Structure

| | | |
|---|----------|--|
| Mother age <20 years at child's birth | Complete | |
| Father age <25 years or >=45 years at child's birth | | Father not registered N=5,834 Father known, age missing N=325 |
| Child's birth order 4 th or greater | Complete | |
| Family size at child's 10 th birthday: | Complete | |
| One, three or four children | | |
| Two children | Complete | |
| Five or more children | Complete | |
| Mother not partnered at time of child's birth | Complete | |

Area Level Socio-economic/demographic Environment

Value of index of socio-economic disadvantage in lowest quintile
Census-year-specific decile allocated for Census Collection District (CCD) of that year according to mapping:

| Census | Birth years |
|--------|-------------|
| 1986 | 1980 – 1987 |
| 1991 | 1988 – 1993 |
| 1996 | 1994 – 1998 |
| 2001 | 1999 – 2001 |

Unable to be classified N=825

Remoteness classification: remote, very remote or unknown

Remoteness value in 2001 allocated by CCD identifier in 2001. ³

Some aggregation of missing values done at postcode level.

Unable to be classified N=68

Missing CCD2001 **and** no postal area aggregation N=353⁴.

Additional missing numbers due to statistics not provided at CCD level:

| | |
|---|-------|
| Percentage of persons of Aboriginal or Torres Strait Islander descent >=25% | N=202 |
| Percentage of persons Australian born <=75% | N=211 |
| Percentage of persons never married >=50% | N=211 |
| Percentage of one parent families >=20% | N=255 |
| Unemployment rate >=5% | N=202 |
| Percentage of persons living same residence as 5 years prior <40% | N=202 |

³ Remoteness classifications only available from 2001 onwards, hence allocation to birth record based on 2001 CCD identifier of residence at birth.

⁴ Postcode information allowed imputation where smaller level CCD identifiers not available for N=23,839

| | |
|---|-------|
| Percentage of persons living same residence as 1 year prior <70% | N=202 |
| Percentage of persons living in semi-detached residences >=10% | N=213 |
| Percentage of persons living in residences that are flats >=5% | N=213 |
| Percentage of persons living in rented residences >=30% | N=216 |
| Percentage of persons who do not speak English well between 1 and 2% | N=202 |
| Has high ethnic heterogeneity | N=202 |
| Total crime - arrest rate per capita in 2002 >=50 per 1,000 residents | |
| Values per Local Government Area (LGA) in 2002 allocated by 2001 LGA | |
| Some missing 2001 LGA mapped by postcode (which is a sub-unit of LGA) | N=5 |
| Has more than minimal level of SEIFA inequality | |
| Distribution of (2001 SEIFA value for CCD 2001 identifier) over postal areas ⁵ | N=8 |

Family Level Sociodemographic Status

| | | |
|---|----------|---|
| Father unknown / not registered at birth | Complete | |
| Father in lower skill occupation or unemployed at time of child's birth | | Father not registered N=5,834 Father known, occupation missing N=6,186 |
| Mother born in low-income country or in Western Australia | | Missing or irreconcilable data N=834 |
| Child of Aboriginal or Torres Strait Islander descent | Complete | |

⁵ Different mechanism to calculate than that used to allocate SEIFA quintile. The inequality approach is necessarily more blunt (all based on 2001 values).

Supplementary Table S2. Adversity exposures for children in development sample excluded or censored due to death. Counts and percentages.

| Adversity exposure | Age child deceased | | | | | |
|--|------------------------|------|----------------------------------|------|--|-------|
| | < 2 days (excluded) | | ≥2 days, <10 years (excluded) | | ≥10 years, before Sz diagnosis or before end of follow-up (censored) | |
| | N= | % | N= | % | N= | 95%CI |
| | 418 | | 840 | | 967 | |
| Discontinuity in Parenting | | | | | | |
| Mother hospitalised ≥ 8 days when child aged 15 days -<1 year | 45 | 10.8 | 34 | 4 | 20 | 2.1 |
| Mother hospitalised ≥ 8 days when child aged 1 -<5 years | 180 | 43.1 | 331 | 39.4 | 244 | 25.2 |
| Mother hospitalised ≥ 8 days when child aged 5 years -<10 years | 79 | 18.9 | 156 | 18.6 | 140 | 14.5 |
| Child hospitalised ≥ 8 days when child aged 15 days -<1 year | | | 149 | 17.7 | 85 | 8.8 |
| Child hospitalised ≥ 8 days when child aged 1 -<5 years | | | 93 | 11.1 | 93 | 9.6 |
| Child hospitalised ≥ 8 days when child aged 5 years -<10 years | | | 37 | 4.4 | 79 | 8.2 |
| Father hospitalised ≥ 8 days when child aged 0 days -<1 year | <5 | <1 | 8 | 1 | 15 | 1.6 |
| Father hospitalised ≥ 8 days when child aged 1 -<5 years | 14 | 3.3 | 30 | 3.6 | 47 | 4.9 |
| Father hospitalised ≥ 8 days when child aged 5 years -<10 years | 17 | 4.1 | 49 | 5.8 | 57 | 5.9 |
| Deceased parents, either or both, before child aged <10 years | <5 | <1 | 27 | 3.2 | 23 | 2.4 |
| Any placement in foster care before child aged <10 years | | | <5 | <1 | 6 | 0.6 |
| Any incarcerations for mother when child aged <10 years | 5 | 1.2 | 16 | 1.9 | 12 | 1.2 |
| Any incarcerations for father when child aged 0 days -<1 year | <5 | <1 | 12 | 1.4 | 13 | 1.3 |
| Any incarcerations for father when child aged 1 -<5 years | 11 | 2.6 | 25 | 3 | 27 | 2.8 |
| Any incarcerations for father when child aged 5 years -<10 years | 10 | 2.4 | 24 | 2.9 | 42 | 4.3 |
| Family Functioning | | | | | | |
| Child protection contact before child aged 10 years: | | | | | | |
| Child or sibling victim of ≥1 substantiated abuse | 15 | 3.6 | 62 | 7.4 | 48 | 5.0 |

| | | | | | | |
|--|-----|------|-----|------|-----|------|
| Child or sibling subject of >=1 unsubstantiated notification, no substantiated cases | 20 | 4.8 | 59 | 7.0 | 85 | 8.8 |
| Neither child nor siblings subject of any notifications | 383 | 91.6 | 719 | 85.6 | 834 | 86.2 |
| Any parental corrective services contact before child aged 10 years | 46 | 11.0 | 128 | 15.2 | 143 | 14.8 |
| Family Structure | | | | | | |
| Mother age <20 years at child's birth | 32 | 7.7 | 75 | 8.9 | 98 | 10.1 |
| Father age <25 years or >=45 years at child's birth | 107 | 25.6 | 237 | 28.2 | 273 | 28.2 |
| Child's birth order 4 th or greater | 64 | 15.3 | 129 | 15.4 | 111 | 11.5 |
| Family size at child's 10 th birthday: | | | | | | |
| One, three or four children | 260 | 62.2 | 526 | 62.6 | 524 | 54.2 |
| Two children | 56 | 13.4 | 123 | 14.6 | 331 | 34.2 |
| Five or more children | 102 | 24.4 | 191 | 22.7 | 112 | 11.6 |
| Mother not partnered at time of child's birth | 368 | 88.0 | 715 | 85.1 | 834 | 86.2 |
| Area Level Socio-economic/demographic Environment^a | | | | | | |
| Value of index of socio-economic disadvantage in lowest quintile | 102 | 24.4 | 227 | 27.0 | 222 | 23.0 |
| Remoteness classification: remote, very remote, or unknown | 62 | 14.8 | 149 | 17.7 | 178 | 18.4 |
| Percentage of persons of Aboriginal or Torres Strait Islander descent >=25% | 16 | 3.8 | 56 | 6.7 | 48 | 5.0 |
| Percentage of persons Australian born <=75% | 302 | 72.4 | 540 | 64.5 | 654 | 67.7 |
| Percentage of persons never married >=50% | 14 | 3.4 | 39 | 4.7 | 40 | 4.1 |
| Percentage of one parent families >=20% | 135 | 32.4 | 251 | 30.0 | 284 | 29.4 |
| Unemployment rate >=5% | 342 | 82.0 | 642 | 76.7 | 749 | 77.5 |
| Percentage of persons living same residence as 5 years prior <40% | 88 | 21.1 | 207 | 24.7 | 236 | 24.4 |
| Percentage of persons living same residence as 1 year prior <70% | 94 | 22.5 | 194 | 23.2 | 220 | 22.8 |
| Percentage of persons living in semi-detached residences >=10% | 107 | 25.7 | 188 | 22.5 | 238 | 24.6 |
| Percentage of persons living in residences that are flats >=5% | 87 | 20.9 | 147 | 17.6 | 167 | 17.3 |
| Percentage of persons living in rented residences >=30% | 131 | 31.4 | 294 | 35.1 | 344 | 35.6 |
| Percentage of persons who do not speak English well between 1 and 2% | 28 | 6.7 | 38 | 4.5 | 45 | 4.7 |
| Total crime - arrest rate in 2002 >=50 per 1,000 residents | 50 | 12.0 | 105 | 12.5 | 132 | 13.7 |
| Has high ethnic heterogeneity | 409 | 98.1 | 810 | 96.8 | 937 | 97.0 |
| Has more than minimal level of SEIFA inequality | 38 | 9.1 | 127 | 15.2 | 138 | 14.3 |
| Family Level Sociodemographic Status | | | | | | |

| | | | | | | |
|---|-----|------|-----|------|-----|------|
| Father unknown / not registered at birth | 39 | 9.3 | 87 | 10.4 | 67 | 6.9 |
| Father in lower skill occupation or unemployed at time of child's birth | 76 | 20.1 | 163 | 21.6 | 186 | 20.7 |
| Mother born in low-income country or in Western Australia | 239 | 57.2 | 530 | 63.1 | 634 | 65.6 |
| Child of Aboriginal or Torres Strait Islander descent | 58 | 13.9 | 135 | 16.1 | 153 | 15.8 |

Supplementary Table S3. Exposure to adversity in validation samples. Number and percentage.

| | Validation sample 1 | | | | Validation sample 2 | | | |
|---|---------------------|----------|------------------|----------|---------------------|----------|------------------|----------|
| | Child with | | Total | | Child with | | Total | |
| | diagnosis of Sz | % | | | diagnosis of Sz | % | | |
| Adversity exposure | N=472 | % | N=128,691 | % | N=492 | % | N=128,691 | % |
| Discontinuity in Parenting | | | | | | | | |
| Mother hospitalised >= 8 days when child aged 15 days <1 year | 14 | 3 | 1188 | 0.9 | 10 | 2 | 1206 | 0.9 |
| Mother hospitalised >= 8 days when child aged 1 <5 years | 138 | 29.2 | 22993 | 17.9 | 142 | 28.9 | 22992 | 17.9 |
| Mother hospitalised >= 8 days when child aged 5 years <10 years | 93 | 19.7 | 12380 | 9.6 | 109 | 22.2 | 12417 | 9.6 |
| Child hospitalised >= 8 days when child aged 15 days <1 year | 32 | 6.8 | 3098 | 2.4 | 26 | 5.3 | 3061 | 2.4 |
| Child hospitalised >= 8 days when child aged 1 <5 years | 36 | 7.6 | 3230 | 2.5 | 37 | 7.5 | 3175 | 2.5 |
| Child hospitalised >= 8 days when child aged 5 years <10 years | 13 | 2.8 | 1496 | 1.2 | 13 | 2.6 | 1580 | 1.2 |
| Father hospitalised >= 8 days when child aged 0 days <1 year | 5 | 1.1 | 727 | 0.6 | 5 | 1 | 706 | 0.5 |
| Father hospitalised >= 8 days when child aged 1 <5 years | 20 | 4.2 | 3695 | 2.9 | 20 | 4.1 | 3637 | 2.8 |
| Father hospitalised >= 8 days when child aged 5 years <10 years | 31 | 6.6 | 4515 | 3.5 | 29 | 5.9 | 4606 | 3.6 |
| Deceased parents, either or both, before child aged <10 years | 8 | 1.7 | 1545 | 1.2 | 11 | 2.2 | 1475 | 1.1 |
| Any placement in foster care before child aged <10 years | 12 | 2.5 | 506 | 0.4 | 8 | 1.6 | 519 | 0.4 |
| Any incarcerations for mother when child aged <10 years | 10 | 2.1 | 478 | 0.4 | 9 | 1.8 | 479 | 0.4 |
| Any incarcerations for father when child aged 0 days <1 year | 12 | 2.5 | 727 | 0.6 | 10 | 2 | 708 | 0.6 |
| Any incarcerations for father when child aged 1 <5 years | 21 | 4.4 | 1735 | 1.3 | 20 | 4.1 | 1699 | 1.3 |
| Any incarcerations for father when child aged 5 years <10 years | 10 | 2.1 | 1805 | 1.4 | 18 | 3.7 | 1697 | 1.3 |
| Family Functioning | | | | | | | | |
| Child protection contact before child aged 10 years: | | | | | | | | |
| Child or sibling victim of >=1 substantiated abuse | 34 | 7.2 | 2556 | 2 | 33 | 6.7 | 2675 | 2.1 |
| Child or sibling subject of >=1 unsubst. notification, no substantiated cases | 47 | 10 | 6074 | 4.7 | 52 | 10.6 | 5916 | 4.6 |
| Any parental corrective services contact before child aged 10 years | 81 | 17.2 | 10813 | 8.4 | 77 | 15.7 | 10780 | 8.4 |
| Family Structure | | | | | | | | |
| Mother age <20 years at child's birth | 44 | 9.3 | 6348 | 4.9 | 49 | 10 | 6287 | 4.9 |
| Father age <25 years or >=45 years at child's birth | 136 | 28.8 | 22922 | 17.8 | 142 | 28.9 | 22670 | 17.6 |
| Child's birth order 4 th or greater | 69 | 14.6 | 11599 | 9 | 57 | 11.6 | 11730 | 9.1 |

| | | | | | | | | | |
|---|-----|------|--------|------|-----|------|--------|------|--|
| Family size at child's 10 th birthday: | | | | | | | | | |
| One, three or four children | 256 | 54.3 | 69757 | 54.2 | 263 | 53.5 | 69676 | 54.1 | |
| Two children | 138 | 29.2 | 49313 | 38.3 | 162 | 32.9 | 49526 | 38.5 | |
| Five or more children | 78 | 16.5 | 9621 | 7.5 | 67 | 13.6 | 9489 | 7.4 | |
| Mother not partnered at time of child's birth | 79 | 16.7 | 10303 | 8 | 78 | 15.9 | 10267 | 8 | |
| Area Level Socio-economic/demographic Environmenta | | | | | | | | | |
| Value of index of socio-economic disadvantage in lowest quintile | 130 | 27.5 | 24248 | 18.8 | 116 | 23.6 | 24233 | 18.8 | |
| Remoteness classification: remote, very remote, or unknown | 66 | 14 | 14412 | 11.2 | 55 | 11.2 | 14274 | 11.1 | |
| Percentage of persons of Aboriginal or Torres Strait Islander descent >=25% | 26 | 5.5 | 2931 | 2.3 | 17 | 3.5 | 2871 | 2.2 | |
| Percentage of persons Australian born <=75% | 326 | 69.1 | 93236 | 72.4 | 361 | 73.4 | 93058 | 72.3 | |
| Percentage of persons never married >=50% | 34 | 7.2 | 4378 | 3.4 | 25 | 5.1 | 4248 | 3.3 | |
| Percentage of one parent families >=20% | 174 | 36.9 | 34353 | 26.7 | 151 | 30.7 | 34340 | 26.7 | |
| Unemployment rate >=5% | 385 | 81.6 | 100283 | 77.9 | 399 | 81.1 | 100060 | 77.8 | |
| Percentage of persons living same residence as 5 years prior <40% | 125 | 26.5 | 28925 | 22.5 | 112 | 22.8 | 28994 | 22.5 | |
| Percentage of persons living same residence as 1 year prior <70% | 117 | 24.8 | 25394 | 19.7 | 114 | 23.2 | 25332 | 19.7 | |
| Percentage of persons living in semi-detached residences >=10% | 148 | 31.4 | 30351 | 23.6 | 142 | 28.9 | 30247 | 23.5 | |
| Percentage of persons living in residences that are flats >=5% | 109 | 23.1 | 22355 | 17.4 | 112 | 22.8 | 22661 | 17.6 | |
| Percentage of persons living in rented residences >=30% | 183 | 38.8 | 37296 | 29 | 164 | 33.3 | 37442 | 29.1 | |
| Percentage of persons who do not speak English well between 1 and 2% | 19 | 4 | 7498 | 5.8 | 70 | 14.2 | 16712 | 13 | |
| Total crime - arrest rate per capita in 2002 >=50 | 69 | 14.6 | 16693 | 13 | 481 | 97.8 | 125416 | 97.5 | |
| Has high ethnic heterogeneity | 457 | 96.8 | 125375 | 97.4 | 59 | 12 | 10301 | 8 | |
| Has more than minimal level of SEIFA inequality | 61 | 12.9 | 10516 | 8.2 | 462 | 93.9 | 116006 | 90.1 | |
| Family Level Sociodemographic Status | | | | | | | | | |
| Father unknown / not registered at birth | 50 | 10.6 | 4301 | 3.3 | 46 | 9.3 | 4337 | 3.4 | |
| Father in lower skill occupation or unemployed at time of child's birth | 105 | 22.2 | 20079 | 15.6 | 83 | 16.9 | 19771 | 15.4 | |
| Mother born in low-income country or in Western Australia | 315 | 66.7 | 76090 | 59.1 | 320 | 65 | 76052 | 59.1 | |
| Child of Aboriginal or Torres Strait Islander descent | 92 | 19.5 | 9188 | 7.1 | 89 | 18.1 | 8916 | 6.9 | |

Supplementary Table S4. Development sample observations excluded from final model parameter estimation due to large influence. N=27 observations, all with schizophrenia diagnosis. Adversity exposure comparison with non-excluded observations.

| Adversity Exposure | Excluded observations | | | | | | Non excluded observations | | | | | |
|--|-----------------------|------|----------------------|------|------------------------|------|---------------------------|---|----------------------|---|------------------------|---|
| | Child with Sz, N=27 | | Child with Sz, N=629 | | Child NO Sz, N=170,932 | | Child with Sz, N=27 | | Child with Sz, N=629 | | Child NO Sz, N=170,932 | |
| | N | % | N | % | N | % | N | % | N | % | N | % |
| Discontinuity in parenting | | | | | | | | | | | | |
| Mother hospitalised >= 8 days when child aged 15 days -<1 year | <5 | <15 | 12 | 1.9 | 1630 | 1 | | | | | | |
| Mother hospitalised >= 8 days when child aged 1 -<5 years | 17 | 63 | 205 | 32.6 | 30937 | 18.1 | | | | | | |
| Mother hospitalised >= 8 days when child aged 5 years -<10 years | 14 | 51.9 | 110 | 17.5 | 16431 | 9.6 | | | | | | |
| Child hospitalised >= 8 days when child aged 15 days -<1 year | 7 | 25.9 | 35 | 5.6 | 4124 | 2.4 | | | | | | |
| Child hospitalised >= 8 days when child aged 1 -<5 years | 5 | 18.5 | 39 | 6.2 | 4175 | 2.4 | | | | | | |
| Child hospitalised >= 8 days when child aged 5 years -<10 years | <5 | <15 | 16 | 2.5 | 1964 | 1.1 | | | | | | |
| Father hospitalised >= 8 days when child aged 0 days -<1 year | <5 | <15 | 9 | 1.4 | 945 | 0.6 | | | | | | |
| Father hospitalised >= 8 days when child aged 1 -<5 years | 8 | 29.6 | 30 | 4.8 | 4818 | 2.8 | | | | | | |
| Father hospitalised >= 8 days when child aged 5 years -<10 years | <5 | <15 | 38 | 6 | 6004 | 3.5 | | | | | | |
| Deceased parents, either or both, before child aged <10 years | <5 | <5 | 11 | 1.7 | 2065 | 1.2 | | | | | | |
| Any placement in foster care before child aged <10 years | 5 | 18.5 | 11 | 1.7 | 692 | 0.4 | | | | | | |
| Any incarcerations for mother when child aged <10 years | <5 | <15 | 8 | 1.3 | 646 | 0.4 | | | | | | |
| Any incarcerations for father when child aged 0 days -<1 year | <5 | <15 | 10 | 1.6 | 882 | 0.5 | | | | | | |
| Any incarcerations for father when child aged 1 -<5 years | 6 | 22.2 | 19 | 3 | 2214 | 1.3 | | | | | | |
| Any incarcerations for father when child aged 5 years -<10 years | 6 | 22.2 | 19 | 3 | 2308 | 1.4 | | | | | | |
| Family functioning | | | | | | | | | | | | |
| Child protection contact before child aged 10 years: | | | | | | | | | | | | |
| Child or sibling victim of >=1 substantiated abuse | 7 | 25.9 | 38 | 6 | 3388 | 2 | | | | | | |
| Child or sibling subject of >=1 unsubstantiated notification, no substantiated cases | 6 | 22.2 | 55 | 8.7 | 8046 | 4.7 | | | | | | |
| Neither child nor siblings subject of any notifications | 14 | 51.9 | 536 | 85.3 | 159498 | 93.3 | | | | | | |
| Any parental corrective services contact before child aged 10 years | 12 | 48 | 78 | 12.4 | 14291 | 8.4 | | | | | | |
| Family structure | | | | | | | | | | | | |
| Mother age <20 years at child's birth | 14 | 51.9 | 53 | 8.4 | 8418 | 4.9 | | | | | | |
| Father age <25 years or >=45 years at child's birth | 8 | 29.6 | 183 | 29.1 | 30284 | 17.7 | | | | | | |
| Child's birth order 4 th or greater | 3 | 11.1 | 78 | 12.4 | 15442 | 9 | | | | | | |
| Family size at child's 10 th birthday: | | | | | | | | | | | | |
| One, three or four children | 19 | 70.4 | 364 | 57.9 | 92540 | 54.1 | | | | | | |
| Two children | <5 | <15 | 180 | 28.6 | 65776 | 38.5 | | | | | | |
| Five or more children | 5 | 18.5 | 85 | 13.5 | 12616 | 7.4 | | | | | | |
| Mother not partnered at time of child's birth | 9 | 33.3 | 96 | 15.3 | 13137 | 7.7 | | | | | | |
| Area Level Socio-economic/demographic Environment | | | | | | | | | | | | |
| Value of index of socio-economic disadvantage in lowest quintile | 14 | 51.9 | 168 | 26.6 | 32494 | 19 | | | | | | |
| Remoteness classification: remote, very remote, or unknown | 9 | 33.3 | 90 | 14.3 | 19159 | 11.2 | | | | | | |

| | | | | | | |
|---|----|------|-----|------|--------|------|
| Percentage of persons of Aboriginal or Torres Strait Islander descent >=25% | <5 | <15 | 27 | 4.3 | 3745 | 2.2 |
| Percentage of persons Australian born <=75% | 14 | 51.9 | 467 | 74 | 124098 | 72.6 |
| Percentage of persons never married >=50% | <5 | <15 | 44 | 7 | 5880 | 3.4 |
| Percentage of one parent families >=20% | 14 | 51.9 | 204 | 32.3 | 45679 | 26.7 |
| Unemployment rate >=5% | 17 | 63 | 512 | 81.1 | 133588 | 78.2 |
| Percentage of persons living same residence as 5 years prior <40% | 10 | 37 | 158 | 25 | 38710 | 22.6 |
| Percentage of persons living same residence as 1 year prior <70% | 10 | 37 | 159 | 25.2 | 34021 | 19.9 |
| Percentage of persons living in semi-detached residences >=10% | 12 | 44.4 | 182 | 28.8 | 40282 | 23.6 |
| Percentage of persons living in residences that are flats >=5% | 6 | 22.2 | 146 | 23.1 | 30214 | 17.7 |
| Percentage of persons living in rented residences >=30% | 17 | 63 | 236 | 37.4 | 50010 | 29.3 |
| Percentage of persons who do not speak English well between 1 and 2% | <5 | <15 | 103 | 16.3 | 21896 | 12.8 |
| Total crime - arrest rate per capita in 2002 >=50 | <5 | <15 | 621 | 98.4 | 167004 | 97.7 |
| Has high ethnic heterogeneity | 7 | 25.9 | 85 | 13.5 | 13696 | 8 |
| Has more than minimal level of SEIFA inequality | 20 | 74.1 | 579 | 91.8 | 154697 | 90.5 |
| Family Level Sociodemographic Status | | | | | | |
| Father unknown / not registered at birth | <5 | <15 | 49 | 7.8 | 5833 | 3.4 |
| Father in lower skill occupation or unemployed at time of child's birth | 11 | 40.7 | 129 | 20.4 | 26701 | 15.6 |
| Mother born in low-income country or in Western Australia | 21 | 77.8 | 413 | 65.5 | 101412 | 59.3 |
| Child of Aboriginal or Torres Strait Islander descent | 14 | 51.9 | 117 | 18.5 | 12156 | 7.1 |

Supplementary Table S5. Parameter estimates derived from development sample including observations with large influence. Comparison with optimum models. Hazard ratios.

| | Influential observations included | Influential observations excluded | |
|--|-----------------------------------|-----------------------------------|------|
| | HR | HR | |
| Model M1 - Exposure to <i>Discontinuity in parenting</i> and risk of schizophrenia. N= 2 excluded | | | |
| Mother hospitalised ≥ 8 days when child aged 1 - <5 years | 1.55 | 1.55 | |
| Mother hospitalised ≥ 8 days when child aged 5 years -<10 years | 1.40 | 1.39 | |
| Child hospitalised ≥ 8 days when child aged 15 days -<1 year | 1.72 | 1.67 | |
| Child hospitalised ≥ 8 days when child aged 1 -<5 years | 1.38 | 1.40 | |
| Child hospitalised ≥ 8 days when child aged 5 years -<10 years | 1.38 | 1.39 | |
| Father hospitalised ≥ 8 days when child aged 1 -<5 years | 1.46 | 1.43 | |
| Father hospitalised ≥ 8 days when child aged 5 years -<10 years | 1.36 | 1.35 | |
| Any placement in foster care before child aged <10 years | 4.95 | 4.71 | |
| Any incarcerations for father when child aged 1 - <5 years | 1.72 | 1.80 | |
| Any incarcerations for father when child aged 5 years - <10 years | 1.65 | 1.55 | |
| Model M2 - Exposure to <i>Family functioning</i> and risk of schizophrenia. N=0 excluded. | | | |
| <i>no observations were excluded from optimum domain model</i> | | | |
| Model M3 - Exposure to <i>Family structure</i> and risk of schizophrenia. N=12 excluded. | | | |
| Father age < 25 years or ≥ 45 years at child's birth | 1.51 | 1.55 | |
| Family size at child's 10 th birthday: | | | |
| One, three or four children | 1.35 | 1.32 | |
| Two children | 2.36 | 2.34 | |
| Five or more children | 1.72 | 1.67 | |
| Mother not partnered at time of child's birth | 1.51 | 1.55 | |
| Model M4 - Exposure to <i>Area level socio-economic/demographic environment</i> and risk of schizophrenia. N=7 excluded | | | |
| Value of index of socio-economic disadvantage in lowest quintile | 1.38 | 1.39 | |
| Percentage of persons never married $\geq 50\%$ | 1.57 | 1.54 | |
| Percentage of one parent families $\geq 20\%$ | 1.13 | 1.14 | |
| Percentage of persons living in residences that are flats $\geq 5\%$ | 1.20 | 1.21 | |
| Percentage of persons who do not speak English well between 1 and 2% | 1.30 | 1.31 | |
| Has high ethnic heterogeneity | 1.40 | 1.40 | |
| Missing CCD information and unable to interpolate domain exposures | 2.53 | 1.00 | |
| Model M5 - Exposure to <i>Family level sociodemographic status</i> and risk of schizophrenia. N=0 excluded | | | |
| <i>no observations were excluded from optimum domain model</i> | | | |
| Global model. N=27 excluded | | | |
| Discontinuity in parenting | calculated as log (HR) of M1 | 2.03 | 1.93 |
| Family functioning | calculated as log (HR) of M2 | 1.92 | 1.90 |
| Family structure | calculated as log (HR) of M3 | 1.58 | 1.67 |

| | | | |
|---|---|------|------|
| Area level socio-economic/demographic environment | | | |
| | calculated as log (HR) of M4 | 1.95 | 1.99 |
| Family level sociodemographic status | calculated as log (HR) of M5 | 1.52 | 1.52 |
| Interaction: Discontinuity in parenting and Area level socio-economic/demographic environment | | | |
| | calculated as {log (HR) of M1} x {log (HR) of M4} | 0.68 | 0.58 |
| Interaction: Family functioning and Family structure | | | |
| | calculated as {log (HR) of M2} x {log (HR) of M3} | 0.67 | 0.64 |

Supplementary Table S6. Elementary validations for models. Robustness of estimated Cox model hazard ratios - comparison with estimated logistic regression odds ratios (Ratios and 95% Confidence intervals). Proportional hazards validity – test for non-zero correlation of Schoenfeld residuals with time (P values).

| | Cox | | |
|---|----------------------|---------------------|----------------------|
| | Proportional Hazards | Logistic Regression | Proportional hazards |
| | HR 95%CI | OR 95%CI | P value |
| Model M1 - Exposure to <i>Discontinuity in parenting</i> and risk of schizophrenia | | | |
| Mother hospitalised >= 8 days when child aged 1 - <5 years | 1.56 (1.32-1.84) | 1.81 (1.52-2.14) | 0.976 |
| Mother hospitalised >= 8 days when child aged 5 years -<10 years | 1.38 (1.13-1.70) | 1.49 (1.21-1.84) | 0.821 |
| Child hospitalised >= 8 days when child aged 15 days -<1 year | 1.67 (1.18-2.37) | 1.66 (1.17-2.36) | 0.083 |
| Child hospitalised >= 8 days when child aged 1 -<5 years | 1.41 (1.00-1.99) | 1.53 (1.08-2.17) | 0.638 |
| Child hospitalised >= 8 days when child aged 5 years -<10 years | 1.39 (0.86-2.23) | 1.46 (0.90-2.36) | 0.907 |
| Father hospitalised >= 8 days when child aged 1 -<5 years | 1.43 (1.02-2.01) | 1.53 (1.08-2.16) | 0.107 |
| Father hospitalised >= 8 days when child aged 5 years -<10 years | 1.34 (0.97-1.87) | 1.37 (0.99-1.91) | 0.121 |
| Any placement in foster care before child aged <10 years | 4.70 (2.75-8.01) | 3.27 (1.90-5.62) | 0.448 |
| Any incarcerations for father when child aged 1 - <5 years | 1.80 (1.08-2.99) | 1.71 (1.02-2.86) | 0.928 |
| Any incarcerations for father when child aged 5 years - <10 years | 1.56 (0.93-2.62) | 1.42 (0.84-2.40) | 0.893 |
| Model M2 - Exposure to <i>Family functioning</i> and risk of schizophrenia | | | |
| Child protection contact before child aged 10 years: | | | |
| Child or sibling subject of >=1 notification, no substantiated abuse | 2.17 (1.64-2.86) | 2.09 (1.58-2.75) | 0.094 |
| Child or sibling victim of >=1 substantiated abuse | 3.51 (2.55-4.82) | 3.56 (2.58-4.91) | 0.075 |
| No notification for any child or sibling within family | Reference | Reference | - |
| Any parental corrective services contact before child aged 10 years | 1.80 (1.42-2.29) | 1.47 (1.16-1.87) | 0.728 |
| Model M3 - Exposure to <i>Family structure</i> and risk of schizophrenia | | | |
| Father age < 25 years or >=45 years at child's birth | 1.56 (1.29-1.88) | 1.61 (1.34-1.95) | 0.678 |
| Family size at child's 10 th birthday: | | | |
| One, three or four children | 1.32 (1.11-1.58) | 1.38 (1.15-1.64) | 0.189 |
| Two children | Reference | Reference | - |
| Five or more children | 2.35 (1.82-3.02) | 2.40 (1.86-3.10) | 0.288 |

| | | | | |
|---|---|------------------|------------------|-------|
| Mother not partnered at time of child's birth | | 1.66 (1.31-2.10) | 1.63 (1.29-2.06) | 0.087 |
| Model M4 - Exposure to Area level socio-economic/demographic environment and risk of schizophrenia | | | | |
| Value of index of socio-economic disadvantage in lowest quintile | | 1.39 (1.13-1.70) | 1.31 (1.07-1.61) | 0.951 |
| Percentage of persons never married >=50% | | 1.53 (1.12-2.11) | 1.58 (1.15-2.17) | 0.377 |
| Percentage of one parent families >=20% | | 1.13 (0.95-1.35) | 1.13 (0.94-1.35) | 0.348 |
| Percentage of persons living in residences that are flats >=5% | | 1.20 (1.00-1.46) | 1.26 (1.04-1.52) | 0.377 |
| Percentage of persons who do not speak English well between 1 and 2% | | 1.31 (1.06-1.61) | 1.33 (1.08-1.64) | 0.245 |
| Has high ethnic heterogeneity | | 1.40 (1.09-1.79) | 1.48 (1.15-1.89) | 0.397 |
| Model M5 - Exposure to Family level sociodemographic status and risk of schizophrenia | | | | |
| Father unknown / not registered at birth | | 1.63 (1.20-2.23) | 1.66 (1.22-2.27) | 0.158 |
| Father in lower skill occupation or unemployed at time of child's birth | | 1.36 (1.12-1.66) | 1.32 (1.09-1.61) | 0.602 |
| Mother born in low-income country or in Western Australia | | 1.18 (1.00-1.39) | 1.17 (1.00-1.39) | 0.326 |
| Child of Aboriginal or Torres Strait Islander descent | | 2.81 (2.27-3.47) | 2.75 (2.23-3.41) | 0.898 |
| Global model | | | | |
| Discontinuity in parenting | calculated as log (HR) of M1 | 1.93 (1.52-2.45) | 2.27 (1.80-2.86) | 0.827 |
| Family functioning | calculated as log (HR) of M2 | 1.90 (1.36-2.66) | 1.55 (1.10-2.17) | 0.913 |
| Family structure | calculated as log (HR) of M3 | 1.66 (1.31-2.11) | 1.78 (1.40-2.25) | 0.692 |
| Area level socio-economic/demographic environment | calculated as log (HR) of M4 | 1.99 (1.46-2.70) | 2.05 (1.51-2.78) | 0.480 |
| Family level sociodemographic status | calculated as log (HR) of M5 | 1.53 (1.26-1.86) | 1.47 (1.21-1.79) | 0.962 |
| Interaction: Discontinuity in parenting and Area level socio-economic/demographic environment | calculated as {log (HR) of M1} x {log (HR) of M4} | 0.58 (0.37-0.90) | 0.55 (0.36-0.85) | 0.546 |
| Interaction: Family functioning and Family structure | calculated as {log (HR) of M2} x {log (HR) of M3} | 0.65 (0.45-0.93) | 0.64 (0.44-0.93) | 0.551 |

Supplementary Table S7. Risk of developing schizophrenia following exposure to adversity across five domains. Alternative Global models. Hazard ratios and 95% Confidence intervals.

| Alternative global model G_1 | | HR | 95%CI |
|--|---|-----------|--------------|
| A1 Discontinuity in parenting | -calculated as log (HR) of M1 | 1.58 | (1.32-1.89) |
| A2 Family functioning | -calculated as log (HR) of M2 | 1.34 | (1.09-1.65) |
| A3 Family structure | -calculated as log (HR) of M3 | 1.54 | (1.24-1.91) |
| A4 Area level socio-economic/demographic environment | -calculated as log (HR) of M4 | 1.60 | (1.23-2.07) |
| A5 Family level sociodemographic status | -calculated as log (HR) of M5 | 1.49 | (1.22-1.81) |
| Alternative global model G_2 | | HR | 95%CI |
| A1 Discontinuity in parenting | -calculated as log (HR) of M1 _i ^a | 2.29 | (1.71-3.08) |
| A2 Family functioning | -calculated as log (HR) of M2 | 2.37 | (1.62-3.48) |
| A3 Family structure | -calculated as log (HR) of M3 | 1.65 | (1.29-2.1) |
| A4 Area level socio-economic/demographic environment | -calculated as log (HR) of M4 _i ^a | 1.78 | (1.39-2.28) |
| A5 Family level sociodemographic status | -calculated as log (HR) of M5 | 1.66 | (1.28-2.17) |
| A1.A2 Interaction | - {log (HR) of M1 _i } x {log (HR) of M2} | 0.69 | (0.47-1.004) |
| A1.A5 Interaction | - {log (HR) of M1 _i } x {log (HR) of M5} | 0.71 | (0.5-1.01) |
| A2.A3 Interaction | - {log (HR) of M2} x {log (HR) of M3} | 0.56 | (0.36-0.87) |

^aAlternative model parameterisations for risk of developing schizophrenia following exposure to adversity across domains A1 – Discontinuity in parenting and A4 – Area level socio-economic/demographic environment. Hazard ratios and 95% Confidence intervals.

| Model M1_i - Exposure to Discontinuity in parenting and risk of schizophrenia | HR | 95%CI |
|--|-----------|--------------|
| Mother hospitalised >= 8 days when child aged 1 - <5 years | 1.67 | 1.39 – 2.02 |
| Mother hospitalised >= 8 days when child aged 5 years -<10 years | 1.72 | 1.31 – 2.25 |
| Child hospitalised >= 8 days when child aged 15 days -<1 year | 1.89 | 1.25 – 2.86 |
| Child hospitalised >= 8 days when child aged 1 -<5 years | 1.65 | 1.12 – 2.43 |
| Child hospitalised >= 8 days when child aged 5 years -<10 years | 1.35 | 0.82 – 2.22 |
| Father hospitalised >= 8 days when child aged 1 -<5 years | 1.30 | 0.90 – 1.87 |
| Father hospitalised >= 8 days when child aged 5 years -<10 years | 1.35 | 0.96 – 1.89 |
| Any placement in foster care before child aged <10 years | 3.11 | 1.58 – 6.11 |
| Any incarcerations for father when child aged 1 - <5 years | 1.43 | 0.80 – 2.54 |
| Any incarcerations for father when child aged 5 years - <10 years | 1.57 | 0.90 – 2.76 |
| Interaction: Mother hospitalised >= 8 days when child aged 1 -<5 years AND Mother hospitalised >= 8 days when child aged 5 -<10 years | 0.64 | 0.43 – 0.97 |
| Interaction: Child hospitalised >= 8 days when child aged 15 days -<1 year AND Child hospitalised >= 8 days when child aged 5 years -<10 years | 0.58 | 0.26 – 1.29 |

| Model M4_i - Exposure to Area level socio-economic/demographic environment and risk of schizophrenia | HR | 95%CI |
|---|-----------|--------------|
| Value of index of socio-economic disadvantage in lowest quintile | 1.57 | 1.28 – 1.94 |
| Percentage of persons never married >=50% | 2.24 | 1.44 – 3.48 |
| Percentage of persons living in residences that are flats >=5% | 1.28 | 1.05 – 1.56 |

| | | |
|---|------|-------------|
| Percentage of persons who do not speak English well between 1 and 2% | 1.45 | 1.15 – 1.83 |
| Has high ethnic heterogeneity | 1.32 | 1.02 – 1.70 |
| Interaction: Percentage of persons never married >=50% | | |
| AND | | |
| Percentage of persons living in residences that are flats >=5% | 0.46 | 0.24 – 0.87 |
| Interaction: Value of index of socio-economic disadvantage in lowest quintile | | |
| AND | | |
| Percentage of persons who do not speak English well between 1 and 2% | 0.58 | 0.34 – 1.01 |

Supplementary Table S8. Alternative Global Cox model parameterisations describing risk of schizophrenia following exposure to adversity. Harrell's Concordance statistics.

| Model describing risk of schizophrenia following exposure to adversity | Harrell's Concordance | |
|---|------------------------------|-------------------------|
| | Validation set 1 | Validation set 2 |
| Global model G_1, as defined in Table S8 | 0.655 | 0.624 |
| Global model G_2, as defined in Table S8 | 0.658 | 0.623 |

Supplementary Figure S1. Distributions of predicted log hazard ratios for five domain models and Global model, derived from development sample.

