

Risk factors for being NEET among young people

Authors: Joe Crowley, Gianfranco Addario, Ekaterina Khriakova, Josefiën Breedvelt
Prepared for: Youth Futures Foundation.



We believe that social research has the power to make life better.

By really understanding the complexity of people's lives and what they think about the issues that affect them, we give the public a powerful and influential role in shaping decisions and services that make a difference to everyone. And as an independent, not for profit organisation we're able to put our time and energy into delivering social research that works for society.

National Centre
for Social Research
35 Northampton Square
London EC1V 0AX
020 7250 1866
www.natcen.ac.uk

A Company Limited by
Guarantee

Registered in England
No.4392418

A Charity registered in England
and Wales (1091768) and
Scotland (SC038454)

This project was carried out
in compliance with ISO20252
designbysoapbox.com

Contents

Executive Summary	5
1. Introduction	9
2. Methods	11
3. Risk factors for NEET status	17
4. The overlap of risk factors	20
5. Which risk factors were most important?	30
6. Developing a risk index for being NEET	33
7. Discussion	37
8. Appendix 1: Additional NEET Risk Index Profiles	40
9. Appendix 2: Measurement of marginalisation	45
10. Appendix 3: Defining NEET status	51
11. Appendix 4: Analysis methods	53
12. Appendix 5: Supporting tables	62
13. Reference List	73

Executive Summary

This report explores the extent and degree of overlap between different forms of marginalisation among young people (aged 13 to 25) in England, and how experiencing multiple types of marginalisation may increase the risk of young people not being in employment, education, or training (NEET).

Data

This study was conducted using Next Steps, a national longitudinal cohort study following a sample of people born in 1989-90 from age 13/14 through to the age of 25. This covers the years 2004 to 2015.

Five domains of marginalisation were explored, covering 19 individual risk factors:

- **Education:** having a special educational need, absence from school, exclusion or expulsion from school, low parental involvement in education, and low academic attainment (by age 25).
- **Family circumstances:** having caring responsibilities after the age of 16, having a parent with a disability, having lived in a single parent household, neither parent having a qualification, having been in care or lived apart from their birth parents in childhood, and having had a child of their own.
- **Health:** having had a limiting disability or health condition and having had a mental health problem.
- **Living standards:** having lived in social housing while growing up, and having lived in a household where either parent was unemployed.
- **Risky behaviour:** engaged in anti-social behaviour, had their parents contacted about their behaviour by either the police or by a social or educational welfare service.

Analysis – main findings

Descriptive analysis was conducted to explore the prevalence of each of the 19 risk factors and their association with a young person's likelihood of being NEET.¹ The vast majority of these young people experienced at least one risk factor. Only 8% of young people experienced none, 14% experienced only one, 52% experienced between two and five risk factors, and 23% experienced over six risk factors. The average number of risk factors experienced by these young people was 4.

The preliminary descriptive results showed that each individual risk factor was associated with an increase in a young person's likelihood of being NEET. They also showed that as the number of risk factors a young person experienced increased, their likelihood of being NEET increased as well. For example, among young people with zero risk factors only 5% were NEET, increasing to 18% among those with 2-5 factors, and 74% for young people with 10 or more. This shows that risk factors for NEET

¹ For this study NEET was measured as having been Not in Education, Employment or Training at the time of the survey conducted when the young person was age 18/19, 19/20, and age 25. There was an interval of four years in data collection, between 2010 and 2015, leading to the gap in the data for NEET status between age 20 and 25.

status did overlap among these young people, and that the overlap of multiple risk factors is associated with an increased risk of NEET status.

Having explored whether each factor was linked to young people's likelihood of being NEET separately. We then investigated how the importance of different risk factors for someone's likelihood of being NEET changed once the influence of the others was controlled for using a logistic regression model. The model identified seven risk factors which were no longer associated with an increased risk of being NEET once the role of the other risk factors was adjusted for. Five of these were indicators related to behaviour, including the three risk factors belonging to the 'risky behaviour' domain, as well as two related to behaviour at school – whether the young person had been expelled/excluded from school or been absent from school. The two other factors which were no longer associated with NEET status were low parental involvement in education and having a disabled parent in the household.

The remaining twelve factors were all associated with a statistically significant increase in risk of NEET status in the logistic regression model. Of these, the six risk factors which were found to be most strongly associated with NEET status, once the others had been controlled for, were not having an academic qualification above level 1, having a limiting disability, having a child before age 21, having a child between age 21 and 25, having poor mental health, and having been identified as having SEN.

Following this, to identify whether certain types of risks factors were more likely to be found together, a cluster analysis was conducted. Six distinct clusters were identified, containing between 2 and 3 risk factors (outlined in the table below).

Clusters of marginalisation

Cluster	Risk factors forming each cluster:
1.	<ul style="list-style-type: none">• Exclusion or expulsion from school,• Parents contacted by the police about young persons' behaviour,• Parents contacted by social or educational services about young persons' behaviour.
2.	<ul style="list-style-type: none">• Having had a child before age 21,• having had a child between age 21 and 25.
3.	<ul style="list-style-type: none">• Lived in a single parent household,• Lived in socially rented accommodation,• Engaged in anti-social behaviours.
4.	<ul style="list-style-type: none">• Special educational needs,• Low academic attainment (does not hold a qualification at Level 2 or higher).
5.	<ul style="list-style-type: none">• Having a parent with a disability,• Having two parents with no qualifications,• Had caring responsibilities after the age of 16.
6.	<ul style="list-style-type: none">• Limiting disability or health condition,• Having experienced a mental health problem.

However, these clusters were fairly weak – meaning that the boundaries between clusters were not strongly defined. In other words, that substantial numbers of people with characteristics in one cluster may also have characteristics from other clusters. This is a challenge to policy making, because this variability in the risk factors affecting young people may make it difficult to target policies effectively.

NEET Risk Index

Finally, to identify which groups of young people in society were more at risk of being NEET, a NEET Risk Index was calculated based on the findings of the logistic regression analysis described above. This showed a number of groups to be at increased risk of being NEET. In particular, this identified young people who had lived in low-income households as children, young people who identified as either bisexual or 'other' when asked about their sexuality, and young people from Black ethnic groups and Pakistani or Bangladeshi backgrounds.

Conclusion

This report found that there was substantial overlap among different forms of marginalisation in the form of risk factors. Each risk factor was individually associated with an increased risk of a young person being NEET. And, having multiple risk factors was linked to a cumulative increase in their likelihood of being NEET. This shows that paying attention to the collective impact of different factors in a young person's life is important to understanding how at risk they may be of poor employment and education outcomes. Using the NEET Risk Index, it is possible to identify broad groups of young people who may be at heightened risk of being NEET, as well identifying the risk faced by individuals based on their unique combinations of risk factors. Although more research will be needed to explore this area, for example, considering the role of timing, when the young person experiences a risk factor and whether this may change its effect. This study highlights the importance of taking a multidimensional approach to understanding labour market exclusion.

1. Introduction

The aim of this report is to explore two areas. First, to what extent young people in England experience different types of marginalisation and how far these experiences overlap. And second, what impact these experiences of marginalisation can have on young people's likelihood of not spending time in employment, education, or training (NEET) between ages 18 and 25.

These are challenging questions to answer, because marginalisation is a multi-faceted and complex phenomenon that can be conceived in different ways. NEET status itself could be considered as a type of marginalisation, while it is considered to be a consequence of marginalisation in this study. Given this, to address these research questions, a clear definition of marginalisation was needed. Our approach to the concept is based on the findings of a rapid evidence assessment (REA) carried out by the National Centre for Social Research for Youth Futures Foundation (Addario et al. 2023) to identify prior definitions of marginalisation. Drawing on these, the REA proposed to define marginalisation as:

“... a state of disadvantage, exclusion, or transgression from a ‘norm’ or typical situation / social model, as defined by wider society. Marginalisation occurs as a result of exclusion by and from social structures and institutions, cultural norms, as well as by day-to-day interactions. Marginalisation often results in feelings of being unheard, not belonging, and of being ‘othered’, as well as adverse economic, health and social outcomes.” (Addario et al. 2023).

In this study, being in employment or education was considered to be the norm (or an expected behaviour and characteristic) which increased people's likelihood of being NEET. The multiple and overlapping elements that could keep someone from achieving employment or education were considered forms of marginalisation.

How far marginalisation in young people's lives can increase their risk of being NEET is an important question for policy makers because it has such a significant impact on young people's lives. Although its prevalence has fallen since its peak in 2012, when 1.25 million of young people aged 16-24 were NEET (Powell, 2021), it continues to affect a substantial minority – around 777,000 in January to March 2023 (ONS, 2023). This matters for people's everyday quality of life. Wages, salaries, and income from self-employment are by far the most common source of income in the UK; accounting for 74% of household income in the financial year 2021-22 (DWP, 2023) and, down the line, ensure the accumulation of resources for retirement.

The effects of experiencing a NEET period can also extend beyond people's immediate economic situation. There is evidence to show that unemployment while young can lead to increased likelihood of depression, future unemployment, and to unhealthy behaviours, such as greater alcohol consumption. Among those who do find work, this is more likely to be low-paid and without training. Continuing in

education has also been shown to be linked to a lower level of mortality and a decreased risk of developing certain health problems (Allen, 2014).

The importance of paid work in society can be seen in the continuous attention national statistics pay to employment rates, with the Office for National Statistics regularly monitoring how many young people aged 16-24 are not in education, employment, or training (2023). However, despite the policy attention, making the transition from school to work can be difficult for young adults. Personal, family, and environmental factors can all make a difference to young people finding paid work, putting some young adults at a disadvantage relative to their peers, and delaying or stopping altogether their integration into the labour market. Likewise, there are inequalities in young people's likelihood of continuing in education, for example, leaving school early and poor educational attainment are important predictors of future educational participation (Maychell et al. 1998).

Research questions

In this report we explore what the factors are that are associated with a condition of disadvantage and marginalisation from the labour market, and how exposure to multiple types of these marginalising factors may combine to increase someone's risk of being NEET further. To do this, the report aims to answer the following research questions:

1. What proportion of young people experience different types of marginalisation, and overlaps of marginalisation, which put them at risk of being NEET up to the age of 25?
2. What are the largest overlaps between different types of marginalisation?
3. Which are the most important overlaps or types of marginalisation for a young person's likelihood of being NEET?
4. Which groups of young people are most likely to experience those forms of marginalisation, or overlaps of marginalisation, that are more strongly associated with NEET status?

2. Methods

Data

Next Steps (formerly known as the Longitudinal Study of Young People in England) was selected for this analysis because it provides a representative sample of young people as they transition from school to work. It collects a rich set of data about their lives, which allows us to unpick not just the relationship between individual risk factors and NEET status, but a wide range of characteristics in combination.

Next Steps is a national longitudinal cohort study following a representative sample of people born in 1989-90 (late Millennials) who were in Year 9 and attending English schools in 2004. The study asks about education and employment, economic circumstances, family life, physical and emotional health and wellbeing, social participation and attitudes. Data collection began in 2004 when cohort members were aged 14 with an achieved sample of 15,770 people out of approximately 21,000. Of these, 7,700 took part in Sweep 8 at age 25, the final timepoint included in our analysis. Fieldwork was conducted face-to-face and includes interviews with both parents and young people.

This project makes use of the longitudinal design of Next Steps (repeated observations from the same individual) by analysing measures captured at multiple points in young people's lives. However, a limitation to the data is that, while respondents were surveyed annually until 2010, the next sweep after this was conducted in 2015-16. As shown in Table 2.1, this means that there is a gap of 5 years between Sweep 7 and 8, which has implications for the interpretation of our analysis – in particular, we do not measure if someone was NEET in that intervening period.

This data is also now a historical dataset, applying to the experiences of young people in 2004-2015. As a result of both developments in policy and society more generally, changes in the different forms of marginalisation identified, their prevalence, and their links to young people's likelihood of being NEET may have taken place in the intervening years. This means that some caution should be used in applying the findings of this study to today's young people. Further considerations on the limitations of the data are discussed in the Discussion and in appendices 2 and 3, as well as potential ways for future research to reproduce this work with a more recent cohort.

Table 2.1. Next Steps data collection

Sweep	1	2	3	4	5	6	7	8
Age	13/14	14/15	15/16	16/17	17/18	18/19	19/20	25
Year	2004	2005	2006	2007	2008	2009	2010	2015

Analysis methods

Survey weights were applied to each respondent in the sample when producing all the descriptive statistics used in this report, and when performing regression analysis. These adjust for survey non-response, making the weighted estimates representative of the cohort of school children from which the sample was drawn – young people in England born in 1989-90 who were in Year 9 and attending English schools in 2004.

Cluster analysis was carried out to identify whether there were any natural groupings among the risk factors included in our analysis (types of marginalisation more likely to occur together). Binary logistic regression was used to identify how the likelihood of someone being NEET changes if they experience a particular risk factor, when controlling for the other factors, and to learn more about the relationships between the different risk factors. These two methods complemented each other. Building upon the finding of the cluster analysis, the regression allowed us to explore whether the groups of risk factors more likely to occur together (found to form clusters), were associated with a more severe risk when they occur together than if they occur separately (interaction effects).

Binary logistic regression was also used to develop a NEET risk index, using the regression coefficients to inform the importance (weights) of the different risk factors that composed the index. A relative risk score was computed for each study participant: this allowed us to observe the distribution of NEET risk across key subgroups of the population, geographical areas, and among young people who belong to groups with protected characteristics (as per Equality Act 2010).

When comparing two or more groups we carried out tests for statistical significance, indicating whether a difference between groups found in our data is likely to be found also in the population of interest (in all young people in England born in 1989-90). All results reported are statistically significant at the 95% level: this means that if we draw 100 samples for the same population, we expect to find these differences between groups 95 times. Consequently, it is unlikely that these are differences that can be found in our data only due to sampling chance.

A summary of the analysis methods used to answer the different research questions of the study is provided in Table 2.2, while additional technical information can be found in Appendix 4: Analysis methods.

Table 2.2. Research questions and analysis method used

Research question 1: What proportion of young people experience overlaps of marginalisation, which put them at risk of being NEET, up to the age of 25?	• Descriptive statistics
Research Question 2: What are the largest overlaps between different types of marginalisation?	• Cluster analysis • Descriptive statistics

Research Question 3: Which are the most important overlaps or types of marginalisation for a young person's likelihood of being NEET?

- Logistic regression
- Descriptive statistics

Research Question 4: Which groups of people are most likely to experience those forms of marginalisation, or overlaps of marginalisation, that are more strongly associated with NEET status?

- Risk index
 - Descriptive statistics
-

Measuring NEET status

The key outcome of interest to this study is young people spending time not in employment, education, or training (NEET). To identify these young people in our analysis we follow the definition proposed by the Office for National Statistics (ONS, 2016), where a person is classified as NEET if they meet none of the following conditions:

- Being in work – this includes any form of paid work, either part-time or full-time.
- Enrolled on an educational course, apprenticeship, government supported employment or training programme, or on a job-related training.
- Or if they are enrolled on an education or training course (as outlined above) which they are waiting to start (or restart).

Using the Next Steps data, NEET status has been defined as closely as possible to this definition, using a similar approach to the Labour Force Survey (LFS), a national survey examining the economic activity in England (ONS, 2013) used to produce national statistics. It checks that young people were not in any form of education, employment, or training in sweeps 6, 7 or 8 using data about each young person's main current activity, as well as additional questions to check they were not in any form of work, education, or training outside of their main activity².

Values across the three Sweeps were then combined, identifying anyone who was NEET at either Sweep 6, 7 or 8, which covers the ages 18/19, 19/20 and 25. The weighted estimates indicate that nearly a fifth (24%) of young people were found to have experienced a period of NEET between the ages of 18 and 25 years old³, with 13% NEET in both Sweeps 6 and 7, rising to 15% in Sweep 8.

While the definition of NEET status is aligned to the approach followed in the Labour Force Survey (LFS) by the Office for National Statistics (ONS), the YFF Youth Voice Consultation – a group of young people who helped shape our analysis and sense-checked our initial findings – identified some limitations with it. For example, the official definition considers NEET people who are unable to work (e.g., those with severe limiting disabilities) as well as people who may be working but are not paid (in particular, people with caring responsibilities, or people volunteering). Despite these limitations, the study follows the LFS

² The approach was not consistent across all the Sweeps of Next Steps, as there were some inconsistencies on how the data was collected and archived. More information about the approach followed in this study can be found in Appendix 2.

³ 18% without applying survey weights.

approach to NEET status to ensure replicability and comparability with other research projects on this topic.

There are also two limitations to how NEET status was operationalised, specific to this study. The first is that it does not differentiate between sustained periods of NEET and being NEET for a short while over the period of one interview. Given this, it may include people who were simply going through a transitory period in their lives when they were not in work or education, but who were not fundamentally excluded from the labour market. As such the “NEET at least once” group will be a combination of those who are genuinely marginalised in the labour market and others who were never fundamentally at risk of poor employment outcomes. The second is that it does not differentiate between being NEET at different ages and that some risk factors may either change in importance or cease to be risk factors entirely at different ages. This limits the ability of this analysis to describe the risk factors for NEET status at each stage of a child’s life, rather it provides a general overview across young adulthood.

Measures of marginalisation: risk factors

To inform the choice of indicators used to measure marginalisation in this project, a narrative literature review was conducted to explore characteristics identified in previous research that were linked to an increased probability of being NEET. The review found multiple potential risk factors linked to being NEET, including, among others, educational pathway (Holmes, Murphy, Mayhew, 2021) and career trajectory, school attainment, caring responsibilities and household factors (Feng et al, 2015), poor health, special educational needs and difficult family circumstances (Powell, 2021), being at risk of offending, and having been in care and living in rural areas (Allen, 2014). Drawing on this, nineteen risk factors for being NEET were identified in the Next Steps dataset, which can be grouped under five domains corresponding to the Equalities and Human Rights Measurement Framework (Equalities and Human Rights Commission, 2017): education, family circumstances, health, justice and personal safety, and living standards. Each indicator was reduced to a binary variable for this analysis, which showed whether the young person experienced that form of marginalisation or not. In the case of justice and personal safety, because our data had limited information on the experience of crime among young people, our measure of this domain focussed on the behaviour of the young person and the name of this domain has been changed to ‘risky behaviour’ to reflect this.

Education domain

- **Special educational needs (SEN):** whether a young person has been identified as “having special educational needs of any kind”.
- **Absence from school:** whether a young person has been “off school for a continuous period of 1 month or more, other than for school holidays” in the 12 months before each survey interview.
- **Exclusions/suspensions from school:** whether a young person has been “suspended from a school for a time” or “expelled from school for good” since the start of the school year.
- **Highest academic qualification below Level 2:** whether a young person’s highest academic qualification was NVQ Level 1 or below by age 25 (or they held no qualifications).
- **Low parental involvement in a young person’s education:** when asked “how involved [they] ... personally feel” in the young person’s “school life” parents who answered, “not at all involved” were

coded as having low involvement in their child's education, and were compared with parents who reported feeling "very", "fairly" or "not very" involved in the child's school life.

Family circumstances domain

- **Caring responsibilities after age 16:** whether the young person had any caring responsibilities from age 17 onwards. This includes regular unpaid care for "ill, disabled or elderly relatives or friends" aged 15 or over, and for "children aged 14 or under", either in the young person's household or elsewhere. Where data was missing for this indicator, we considered as having caring responsibilities young people who reported they had a child or a disabled parent. If data was still missing, this measure was taken from a comparable set of questions about pre-16 caring responsibilities (see **Appendix 2** for full details).
- **No parental qualifications:** if both parents held no qualification at the latest Sweep for which data is available (either Sweep 1 or 2 - when the young person was aged 13/14 or 14/15).
- **Parental disability:** whether either parent had a "longstanding illness, disability, or infirmity which limits their day-to-day activities".
- **Single parent household:** whether the young person lived in a single parent household at the time of the survey interview.
- **Having a child:** this was included in our analysis as two separate risk factors. These show whether by age 25 the young person has a child living in their household aged 0 – 4, meaning they were born when the young person was aged 21 to 25, and whether they have a child aged 5 – 11 (meaning they were born before age 21). This relates only to children where the young person is the parent. The age of the respondents is indicative, as study participants were aged between 24 and 26 at the time of the Sweep 8 interview.
- **Having been in care or spent time living apart from birth parents:** this is a combined indicator based on: (1) whether the young person was not living with their natural parents at the time of the survey interview and (2) whether the young person had ever been in care. Being in care could mean having lived with foster parents, in a young people's home, or in local authority care, as well as having been placed for adoption.

Health domain

- **Disability status:** whether the young person has ever had a longstanding illness, disability, or infirmity which limits their activity.
- **Mental health:** whether the young person had symptoms of depression in the latest available data for each young person (asked of most at age 25). This was measured with the General Health Questionnaire-12 survey instrument (GHQ), which is used to screen people for mental disorders. Young people who scored 4 or higher out of 12 on the GHQ were classified as experiencing depression (James et al., 2013).

Living standards domain

- **Social housing:** whether the young person's household lived in accommodation "rented from a Council or New Town" or "rented from a Housing Association" at the time of the survey interview. This was included as an indicator of potential material deprivation.

-
- **Parental unemployment:** whether either of the young person's parents were unemployed or looking for a job at the time of the survey interview.

Risky behaviour

- **Anti-social behaviour:** whether the young person engaged in any of the following activities: graffitiing on walls, vandalising public property or shoplifting.
- **Contact from social or educational welfare services:** whether the main parent had been contacted by social services, educational welfare services or any other similar services about the young person's behaviour in the 12 months before the survey interview.
- **Contact from the police:** whether the police had been in touch with the main parent about something the young person did in the 12 months before the survey interview.

In the two indicators above the "main parent" is the parent or person in the young person's household who is most involved in their education.

Limitations of the data for risk factors

Although the list above covers many areas of people's lives, there are some forms of marginalisation of relevance to a young person's likelihood of being NEET that we could not find in the Next Steps data. In particular, the available data about alcohol consumption and drug use were not included because the way these were collected changed over time and the most in depth measures were collected only at the most recent time points. The other factor which could not be included was educational attainment at different timepoints, rather than overall attainment by age 25 (although a proposal for addressing this is described in the Discussion, under potential areas for future research).

In some cases, the measures we included can only give a partial measure of that risk factor. As described in the list above, living in social housing is included as an indicator of material deprivation, but it will likely fail to include some people who were living in poverty and perhaps include others who were actually not materially deprived. Other data sources would have provided a more complete set of data on poverty and allowed it to be measured more completely, however, this dataset was selected because of its breadth of coverage and its longitudinal dimension, meaning this research can look across multiple domains of marginalisation to understand their overlaps, over a decade. It is important to acknowledge though that this breadth of coverage has required a lower level of detail in some domains.

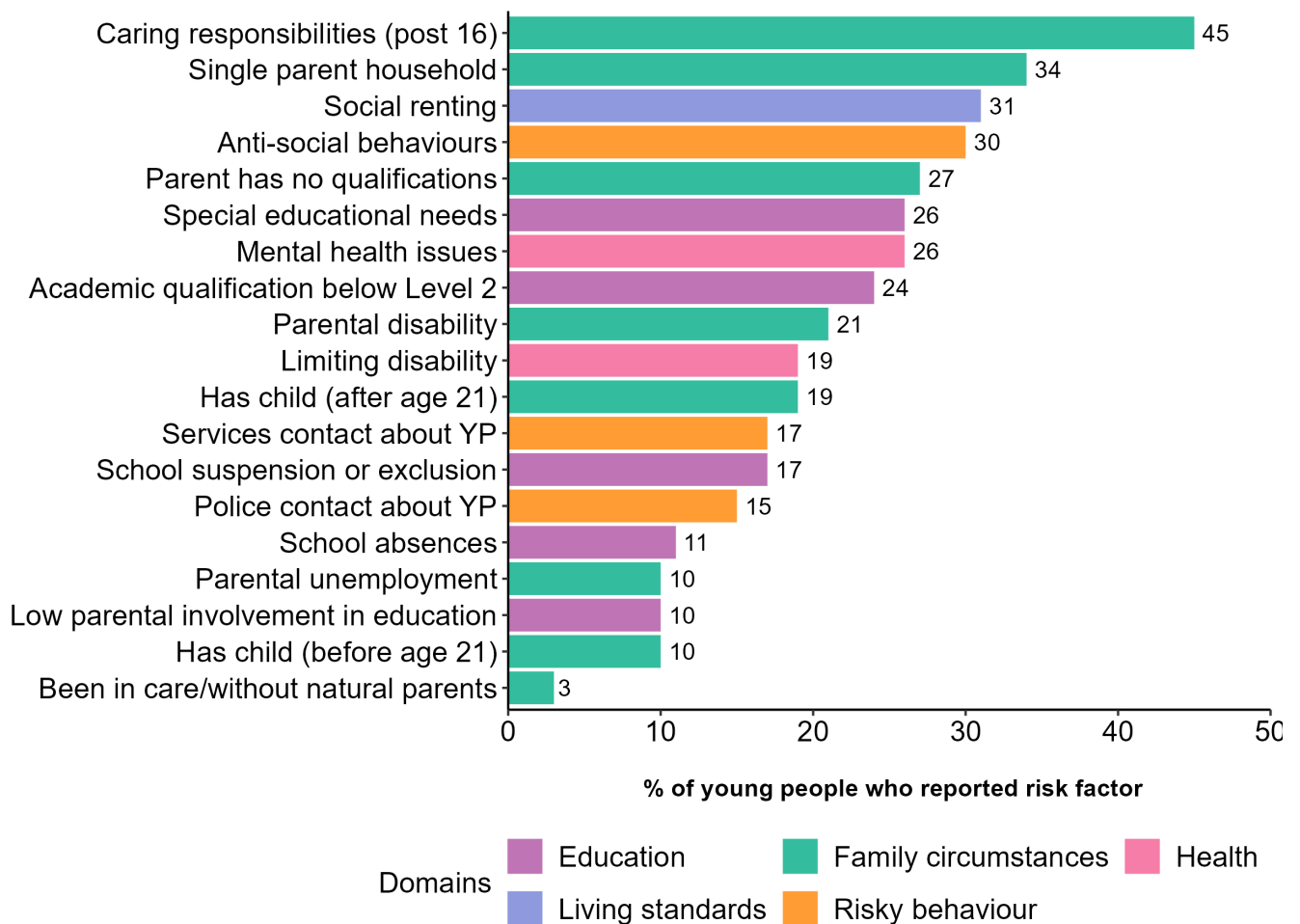
Risk factor timepoints

Most of the risk factors were created using indicators from interviews carried out between ages 13/14 and 16/17, and therefore precede the outcome of interest (NEET status) which was measured at age 18 and after. However, a small number of risk factors were measured using data collected at other timepoints, including age 25. The rationale behind the timepoints selected was to measure risk factors at the earliest possible age points, and consistently over time wherever possible. More information about how many timepoints were used for the different risk factors, and when the information was collected, can be found in Appendix 2.

3. Risk factors for NEET status

As a first step, Figure 3.1 below gives an overview of how common each different form of marginalisation was⁴. It shows that many of the risk factors identified were very common, with the most frequently reported – caring responsibilities after the age of 16 – affecting nearly half of young people (45%) and only 8% of young people experiencing zero risk factors. However, the prevalence of risk factors varies widely, with several that affect only around 1 in 10 young people and the least common, having been in care or lived apart from their natural parents, experienced by only 3% of young people.

Figure 3.1. Proportion of young people who experienced each risk factor



Source: Next Steps Sweep 8 (age 25). Unweighted base size: 7,228 – 7,707.

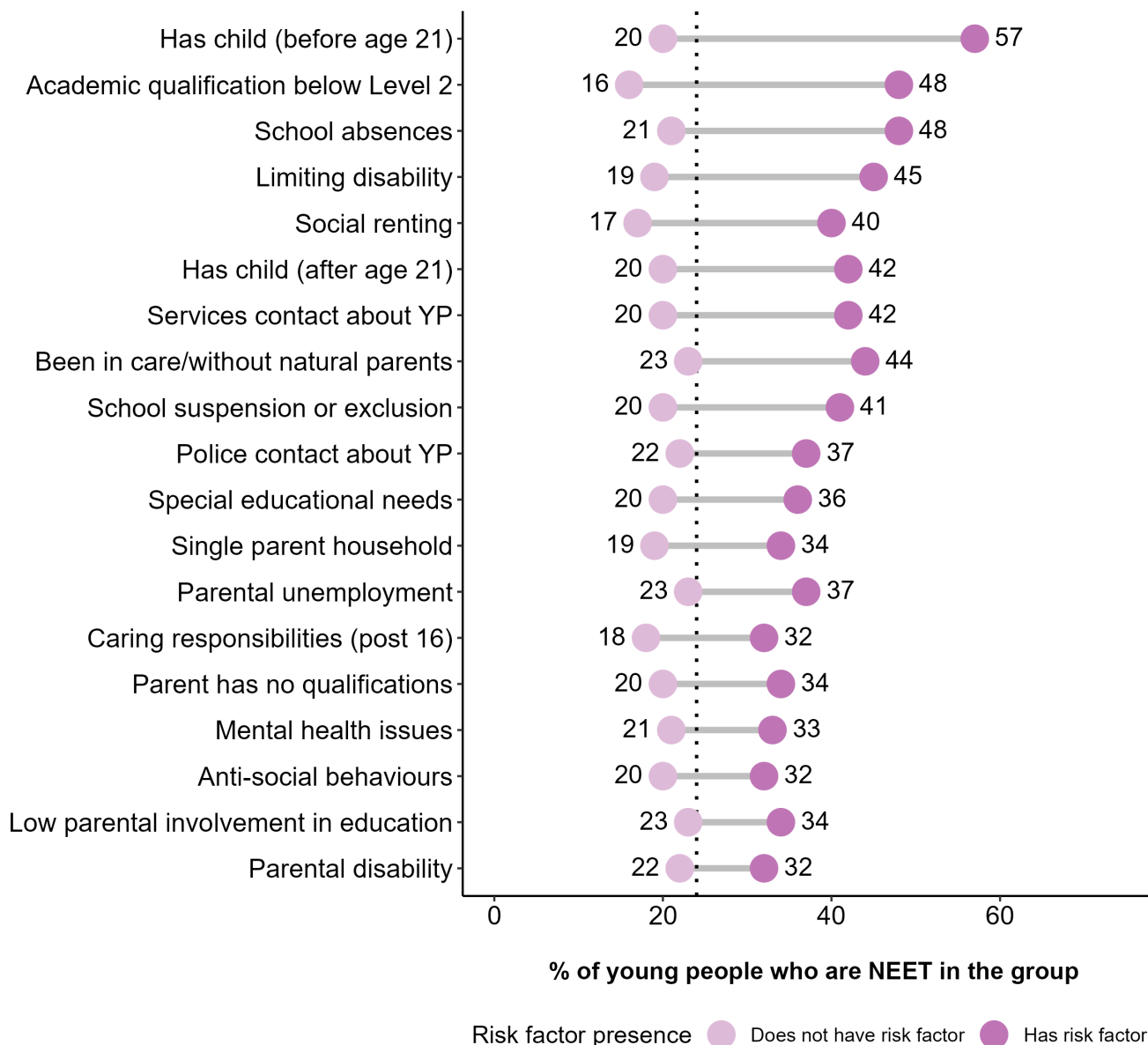
⁴ A limitation to how marginalisation has been measured in this study is that it is likely to underestimate how common the risk factors for NEET status are. This is because, for young people where a question is asked at multiple Sweeps and a person responds to the survey at only one Sweep, we have used the data from the one Sweep where they responded. However, people who were missing at one timepoint and did not experience the risk factor in the Sweeps where data was collected, may have experienced it in the Sweep where they did not respond. The result is that in practice the proportion of people who experienced risk factors measured over multiple waves is probably slightly underestimated.

That fewer than 1 in 10 young people have experienced zero risk factors should be seen in light of the fact that they are defined quite inclusively. For instance, people need to have only lived in social housing at one timepoint from ages 13/14 to 19/20 to be in the “social renting” risk factor. They do not necessarily indicate a long-term experience, just that it has occurred at least once for that young person. Considering that the risk factors are relatively common, the presence of each individual risk factor may not imply that a young person is ‘marginalised’. However, their compound impact on a young person can be considerable, as will be shown later (in Chapter 4).

How much more likely were young people with each risk factor to be NEET?

Figure 3.2 below shows that for all risk factors, the likelihood of being NEET (at some point between age 18 and 25) increases among those young people that have experienced them. Although the size of the difference in the percentage of NEET individuals among those who had and did not have specific risk factors varied substantially.

Figure 3.2 Proportion of young people who have been NEET by risk factor



Source: *Next Steps Sweep 8 (age 25)*. Unweighted base size: 7,199 – 421.

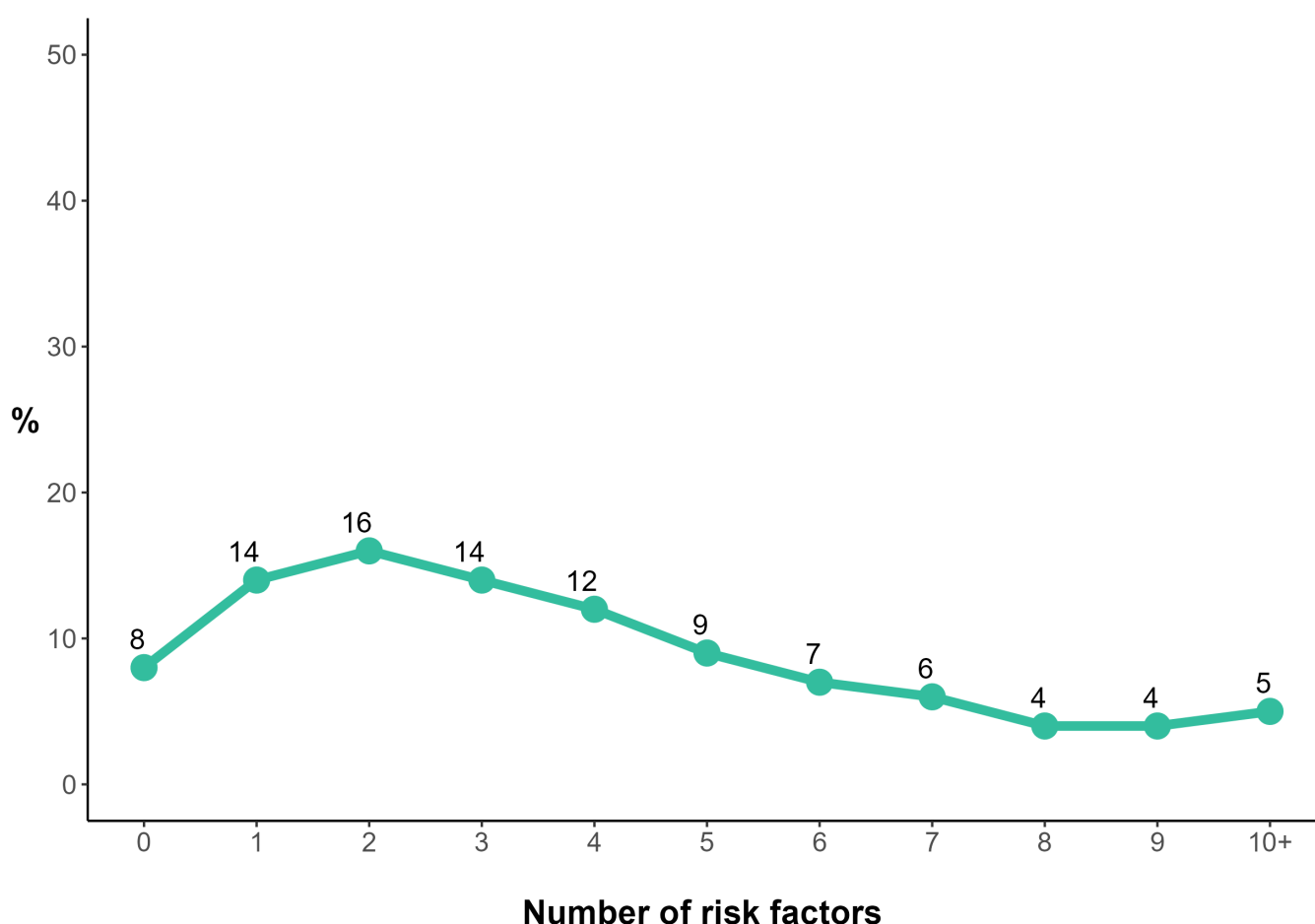
In the chart the pale purple dot on the left shows the percentage of young people who have been NEET among those without the risk factor, and the darker dot shows the percentage among young people who did experience that risk factor. The grey line illustrates the size of the difference between the two, and the risk factors have been ordered by the size of that difference – with the largest change at the top. In other words, the group at the top has the greatest difference in the percentage of young people who were NEET between those with and without the risk factor, and the group at the bottom has the smallest difference. The black dotted vertical line shows the average percentage of young people who are NEET in the population of interest.

This shows that becoming a parent at an early age (before 21) was strongly linked with being NEET, given that it has the greatest difference in NEET status. Among young people who had a child before the age of 21, 57% had been NEET between age 18 and 25, compared to only 20% among those who did not have a child before age 21. Conversely, at the bottom of the chart the difference in the percentage of NEET individuals was much smaller. For young people who had at least one disabled parent, 33% had been NEET compared to 22% among those without a disabled parent.

4. The overlap of risk factors

In the previous section we reported how common each individual risk factor was. Their prevalence suggested that many young people would experience multiple types of marginalisation. Figure 4.1 illustrates how common this experience was. It shows what percentage of young people had different numbers of risk factors, with most clustered at the lower end of the scale. Eight percent of young people had zero risk factors, 14% had only one risk factor, the majority (52%) fell between two and five on the scale, and the remaining quarter (23%) were in the upper end - with 6 or more risk factors. The average number a young person experienced was 4.

Figure 4.1. Prevalence of multiple risk factors

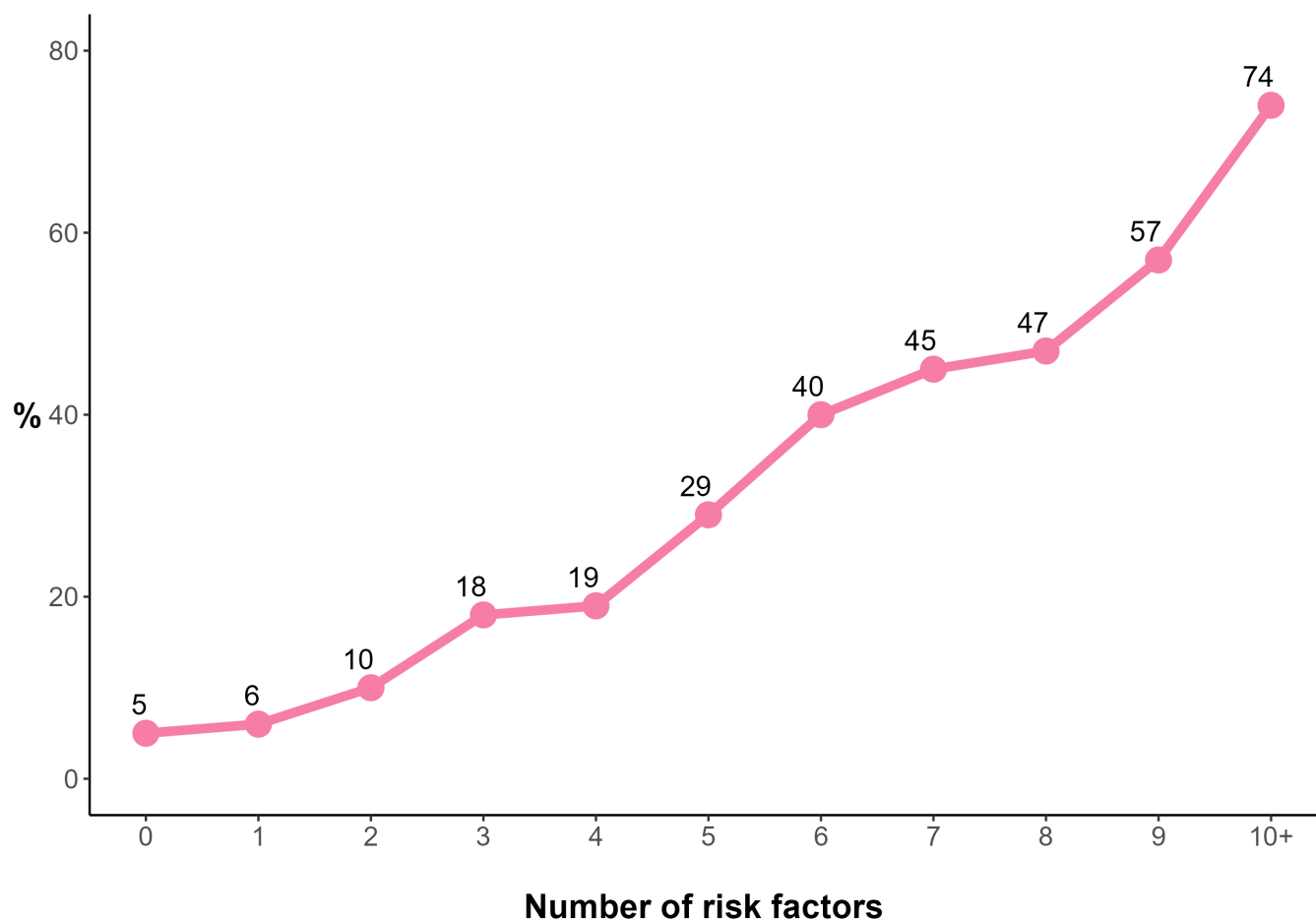


Source: Next Steps Sweep 8 (age 25). Unweighted base size: 7,091.

As the number of risk factors a young person experienced rose, their likelihood of being NEET also increased substantially. Figure 4.2 presents the percentage of young people who had been NEET from age 18 to 25 depending on how many risk factors they reported. It shows a steady increase in NEET status as the number of risk factors rises. While only 5% of young people with zero or one risk factors

were NEET, this percentage increased to 18% among young people with 2 to 5 risk factors, and for those with the ten or more risk factors it reaches 74%.

Figure 4.2 Prevalence of NEET status, by number of risk factors



Source: Next Steps Sweep 8 (age 25). Unweighted base size: 7,082.

Were some risk factors more likely to occur together?

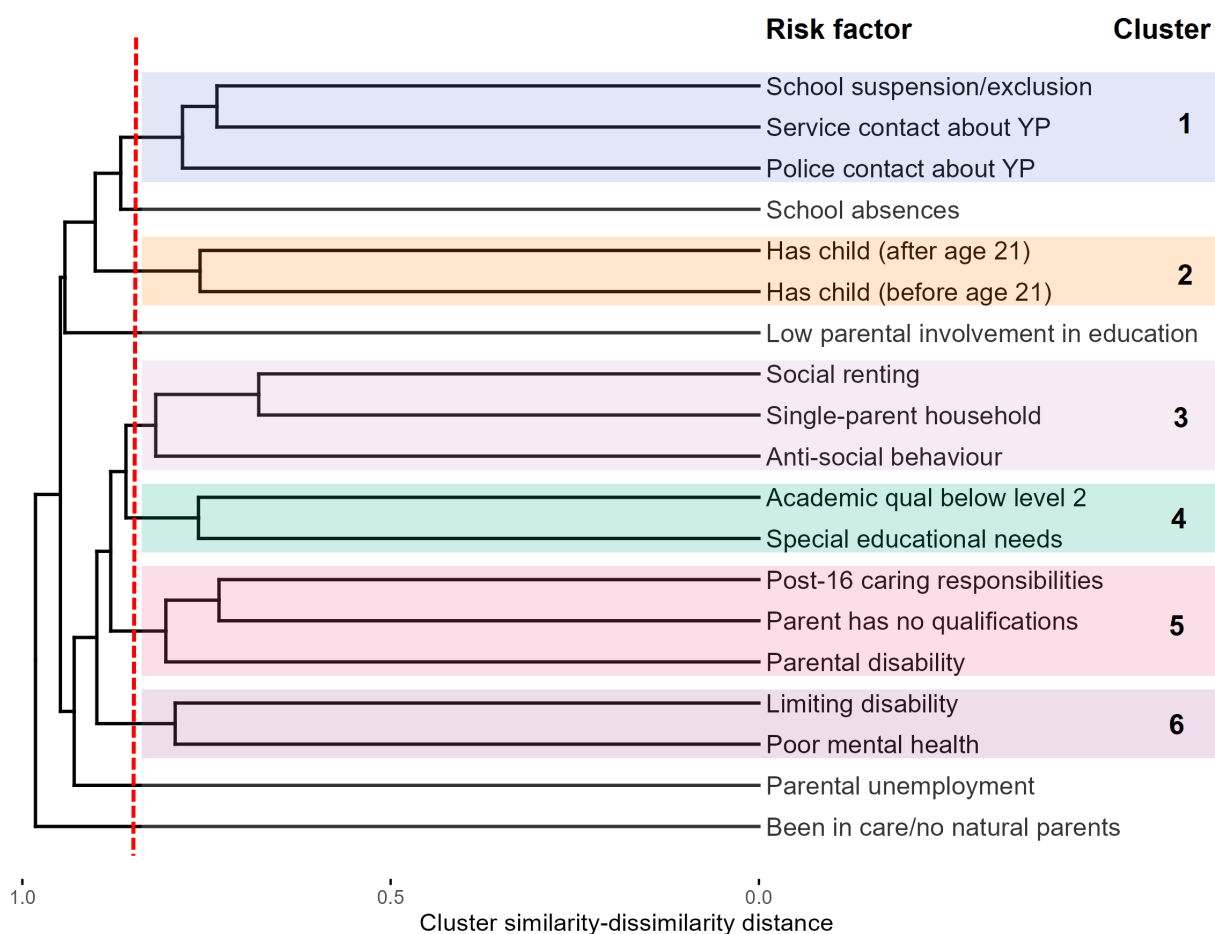
To explore if there were any natural groupings formed by different types of marginalisation, a technique called cluster analysis was applied. Cluster analysis identifies similarities within a dataset and groups more similar cases together. In this case, we clustered risk factors by looking at how likely they were to occur together. Through this analysis, we identified six groups of different types of marginalisation, each of which formed a distinct cluster, although these were relatively weak and grouped together quite late in the clustering process. The clusters identified are illustrated in the dendrogram below (Figure 4.3). The red line represents the cut-off point we have chosen for what can be considered to be a cluster and what cannot: those risk factors whose lines join together before the red line are those that formed clusters together. For example, the first three risk factors at the top of the graph form a cluster, which seems to be related to behavioural issues, but the fourth risk factor – absence from school – does not join it, because it clusters with these three at a later stage and after the red line. This suggests that it is still related to these other factors, which makes sense conceptually, but its association with them is weaker.

It should be noted that experiencing the characteristics in one cluster does not imply that someone will not also experience risk factors from another cluster. These clusters are not mutually exclusive, and only indicate risk factors that are more likely to be experienced together. A person can experience risk factors from **multiple** different clusters.

For each cluster of risk factors identified we report how common the overlap of those risk factors was among young people and how far a young person having that overlapping set of characteristics experienced an increased likelihood of being NEET. In particular, we assessed through a series of logistic regression models whether the **combined risk** emerging from the cluster (the risk of facing NEET status emerging from experiencing one or more risk factors within the same cluster) was different than the **additive risk** (the sum of the risk associated with the different risk factors in the cluster).

For most clusters we did not find evidence suggesting that the combined risk is different from the additive risk. This means that while individual risk factors increased one's likelihood of being NEET, the risk of being NEET associated with each item did not change depending on the other variables. The exception was in Cluster 2, which included special educational needs and low academic attainment. In this case, while there was an increased risk of NEET associated with both factors together, the compound effect of these risk factors was slightly lower than a simple sum of their individual effects.

Figure 4.3 Cluster analysis dendrogram

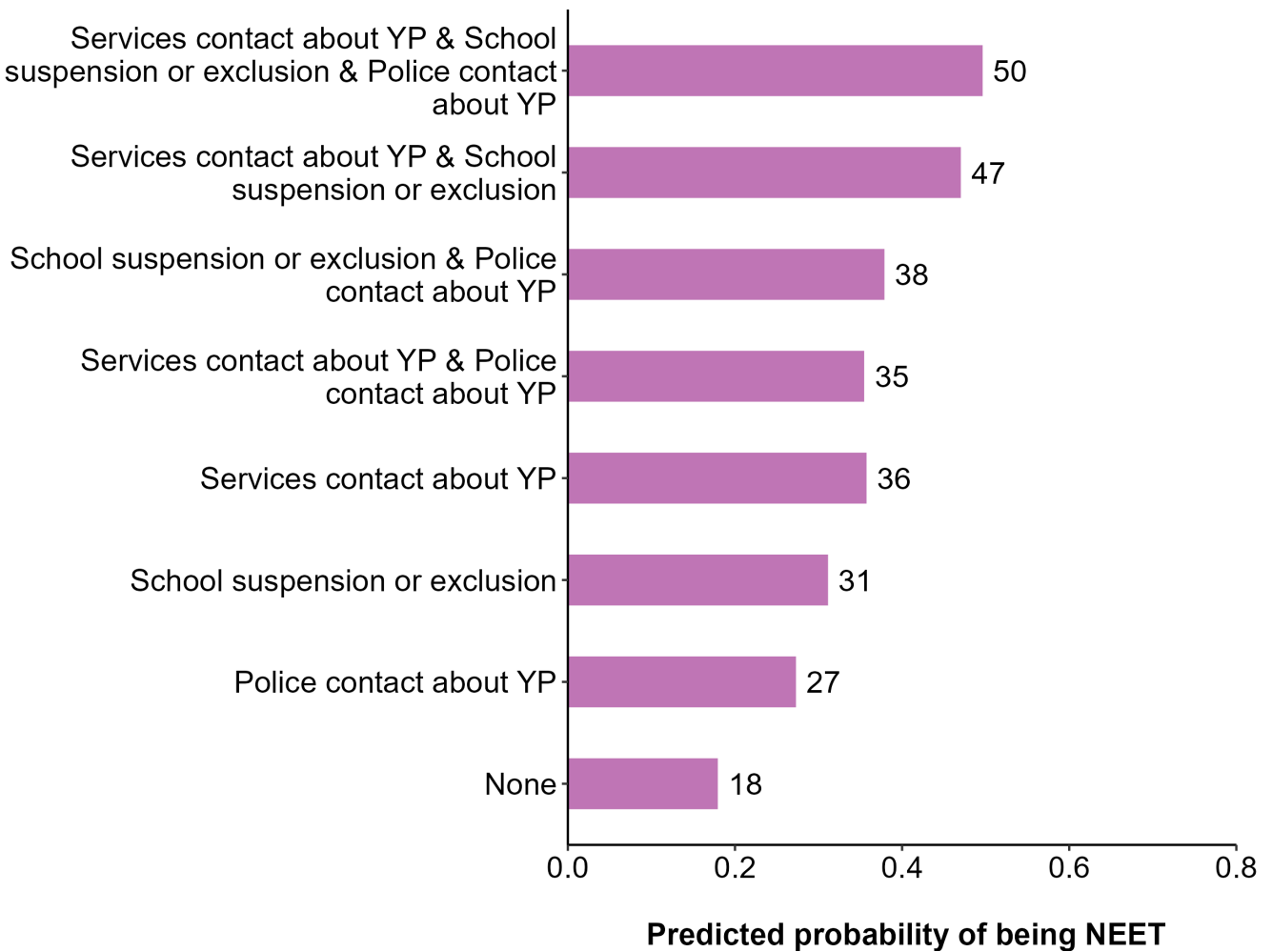


Source: Next Steps Sweep 8 (age 25). Base size: 7,091.

Cluster 1

Cluster 1 seemed to be defined by behaviour issues, including having been expelled or excluded from school, having had the police contact a child's parents about their behaviour, and having had social or education services contact the parents about the child's behaviour. Around one third (31%) of young people had been affected by at least one of these risk factors, 13% had been affected by up to two of them, and 4% by the overlap of all three. As in other clusters those young people who experienced multiple risk factors from this cluster were at greater risk of being NEET due to the additive risk associated with the different factors, but we did not find evidence of a combined effect (Figure 4.4 on the following page).

Figure 4.4 Predicted probability of being NEET for each group in Cluster 1



Source: Next Steps Sweep 8 (age 25). Unweighted base size: 7,328.

Who is at greater risk of being in Cluster 1?

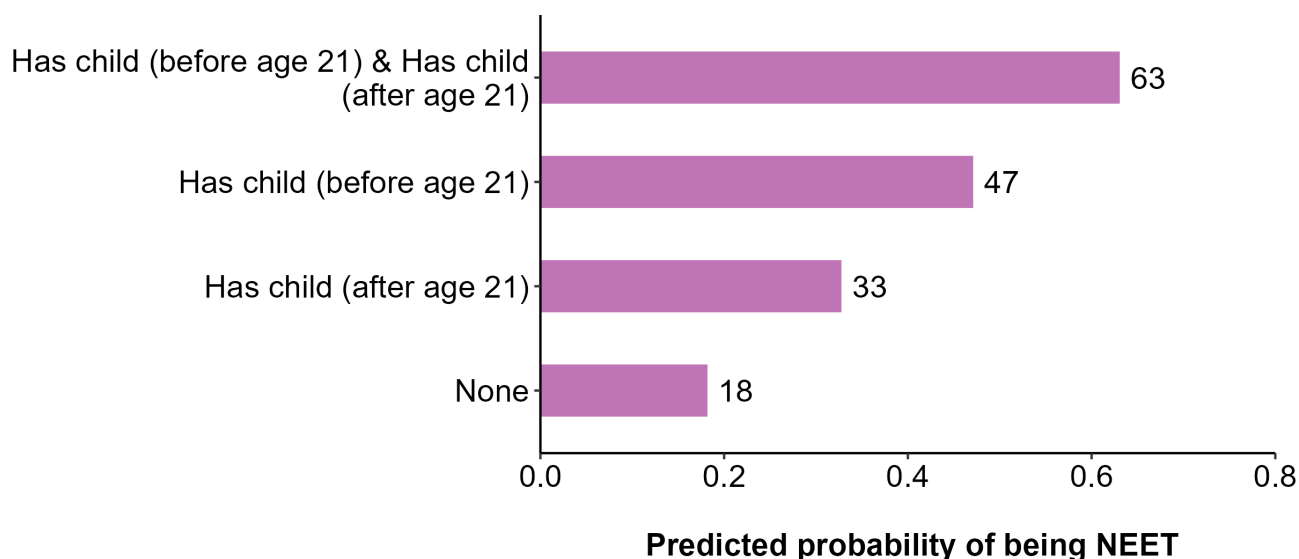
Young people from mixed and white ethnic backgrounds were more likely to experience the overlap of all three variables from Cluster 1, affecting 20% and 14% respectively, than either Asian or black ethnic groups, where it was experienced by 6% and 9% respectively. The overlap was also more common among young people who identified as bisexual – 21% of whom reported all Cluster 1 risk factors, while

the figure was 16% among those of gay or lesbian sexual orientation and 13% for heterosexual young people. Men were more likely to experience the overlap than women – 19% of men reported all of the risk factors from Cluster 1 compared to 10% among women. Among young people who grew up in low-income households the overlap was more common than among young people from more affluent households: 18% reported the overlap in the former group and 11% in the latter group. When looking at regional distribution, combination of risk factors was less common among young people in London (10%), North West (11%) and West Midlands (11%) and more common in North East of England with 19% of young people from the region being affected by the overlap of risk factors from Cluster 1.

Cluster 2

The second cluster identified an association between having a child between ages 21 and 25 and having had a child before the age of 21. This was a relatively small cluster, with 77% of young people still having no children by age 25, and only 6% reported having had a child both before age 21 and after (Figure 4.5). Notably, this is more common than only having had a child before age 21, reported by only 4%. It shows that experiencing multiple risk factors from this cluster was associated with a greater risk of being NEET, and that of the two, having had a child before age 21 was a more important risk factor than having a child after age 21.

Figure 4.5 Predicted probability of being NEET for each group in Cluster 2



Source: Next Steps Sweep 8 (age 25). Unweighted base size: 7,690.

Who is at greater risk of being in Cluster 2?

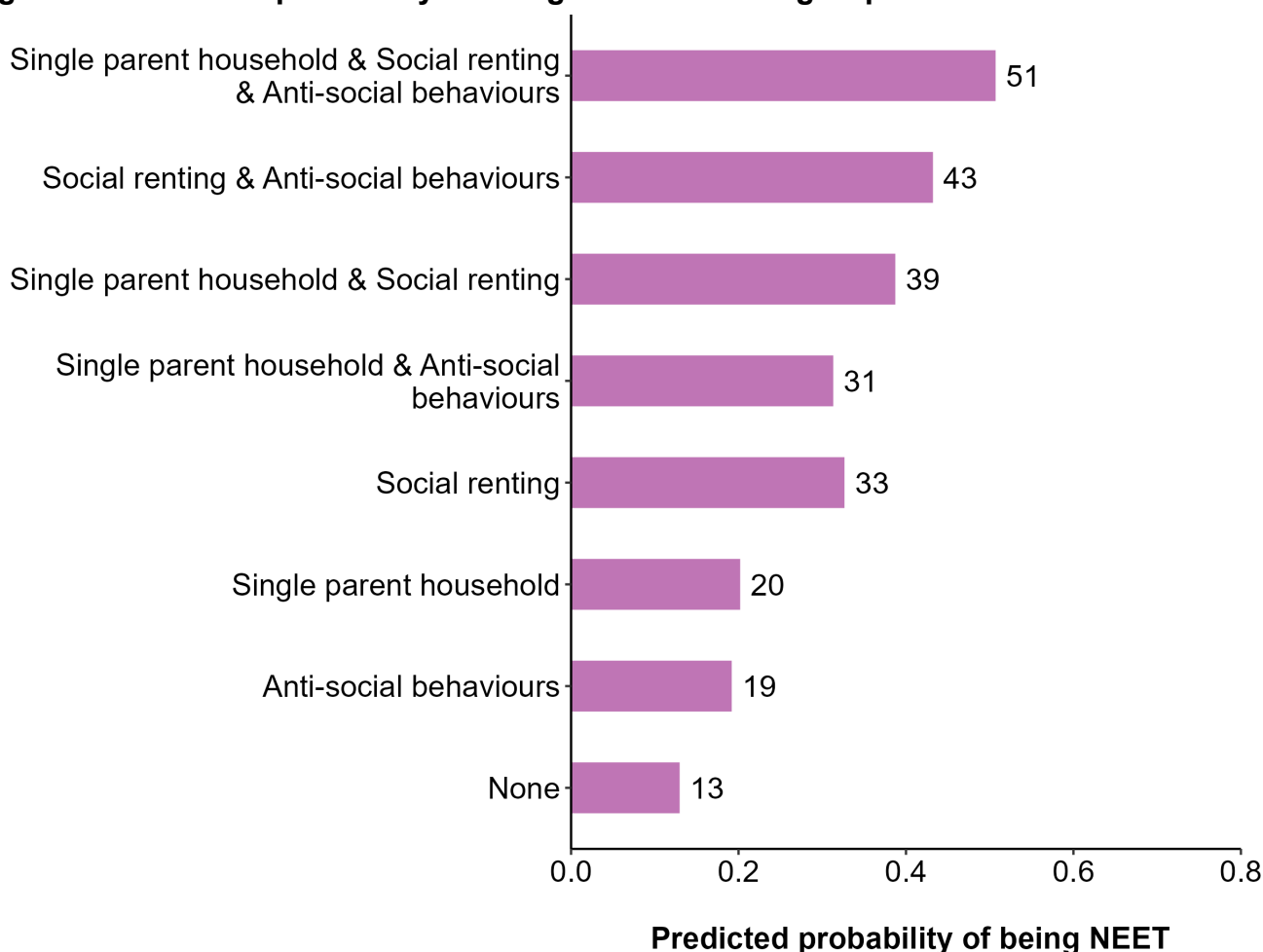
Women were three times more likely to be in Cluster 2 than men – 10% of women had both had a child before age 21 and between the ages of 21 and 25 compared to 3% among men. Some ethnic groups were also at greater risk of being in this cluster, young people of mixed and white ethnicity had the greatest likelihood, experienced by 7% in both cases. This was followed by young people of Pakistani & Bangladeshi ethnicity (4%) and black ethnic groups (4%), but very rare among young people of Indian

ethnicity and other Asian ethnicities (less than 1%). It was also twice as common among young people who lived in low-income households than among those in more affluent households, experienced by 6% compared to 3%. Experiencing an overlap of risk factors was rare among young people from London with only 2% reporting it, and more common in North East (10%), East Midlands (9%) and South West (8%) regions of England. Additionally, bisexual young people were more likely to have a child before age 21, with 9% of them reporting it compared to 6% among heterosexual and zero percent among gay or lesbian young people.

Cluster 3

The third cluster identified a group of variables which included having lived in a single parent household, having lived in socially rented accommodation, and having engaged in anti-social behaviour. Anti-social behaviour included having done any of the following: shoplifting, graffitiing, or vandalising public property. A majority (60%) of young people experienced at least one of these risk factors, 27% experienced at least two, and 7% fell into the overlap of all three. As shown in Figure 4.6, the trend within this cluster was towards a greater likelihood of being NEET for young people who experienced more than one risk factor; this can be explained with the additive risk associated with the number of risk factors experienced.

Figure 4.6 Predicted probability of being NEET for each group in Cluster 3



Source: Next Steps Sweep 8 (age 25). Unweighted base size: 7,601.

In Cluster 3 it also seems the risk predicted by having lived in socially rented accommodation was similar to the effect of both having lived in a single parent household and having engaged in anti-social behaviours. This indicates that, within this cluster, social renting is the risk factor with the strongest association with a young person's likelihood of being NEET. However, it is important to note that while the risk effect of social renting is predominant in this cluster, once we control for other risk factors (those outside this cluster), it seems to lose importance. See the discussion in Chapter 4 and the full regression output in appendices 5.

Who is at greater risk of being in Cluster 3?

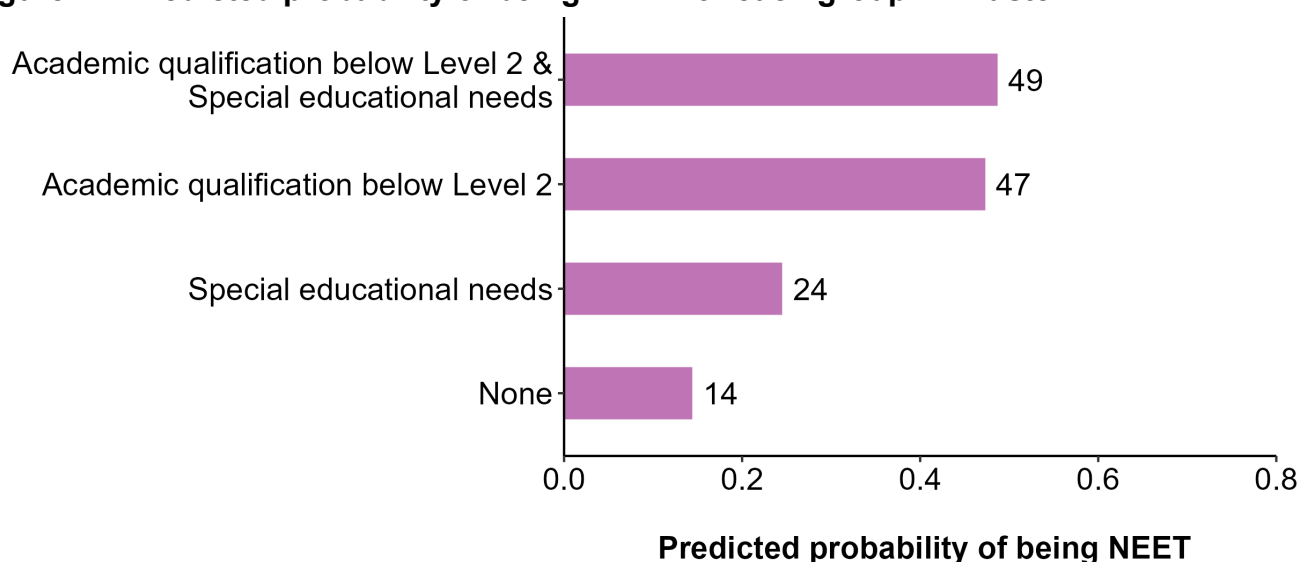
The overlap of any risk factors from Cluster 3 was much more common among young people from black and mixed ethnic group: about a half of them experienced the combination of all three factors, compared to 26% for young people of white ethnicity, 18% for those of Pakistani or Bangladeshi ethnicity and 7% among young people of Indian ethnicity. The overlap of risk factors was also related to income: among young people who grew up in low-income households 46% reported all three factors simultaneously, while among young people from more affluent households only 19% experienced this overlap.

Additionally, bisexual, and gay or lesbian young people were more likely to experience the overlap of risk factors: about a third (33% and 32%) of them reported it compared to 27% for heterosexual young people. When looking at the regional distribution, the overlap was most common in London and East Midlands, where 32% and 30% of young people reported multiple risk factors, and least common in Yorkshire and the Humber regions, where this figure was 22%.

Cluster 4

The fourth cluster identified a group of variables which included young people identified as having SEN and those who had not achieved a Level 2 academic qualification by age 25 (see Figure 4.7 on the following page). These characteristics were fairly common, affecting 1 in 4 young people, and the overlap of these two risk factors was experienced by 12%, meaning it was one of the most common clusters. In this case, although there was an increased risk of NEET associated with both factors together, the compound effect of these risk factors is slightly lower than a simple sum of their individual effects. This can be seen in the fact there is a very similar predicted probability of being NEET (ranging from 0 to 1, with 1 meaning that NEET status is always expected to happen) for the group who only lacked a Level 2 qualification (0.47) compared to that which also had SEN (0.49). From these findings it is not possible to say why this is the case. It suggests that the importance of qualifications for the labour market success of young people with SEN is lower than for young people generally. It may also be that young people with SEN who do not have a Level 2 qualification may have different aspirations in the labour market, compared to young people with SEN and a higher level of qualification, potentially making finding a job easier for them. Additional research would be needed to understand *why* this pattern is seen among young people with SEN who do not hold a Level 2 qualification.

Figure 4.7 Predicted probability of being NEET for each group in Cluster 4



Source: Next Steps Sweep 8 (age 25). Unweighted base size: 7,649.

Who is at greater risk of being in Cluster 4?

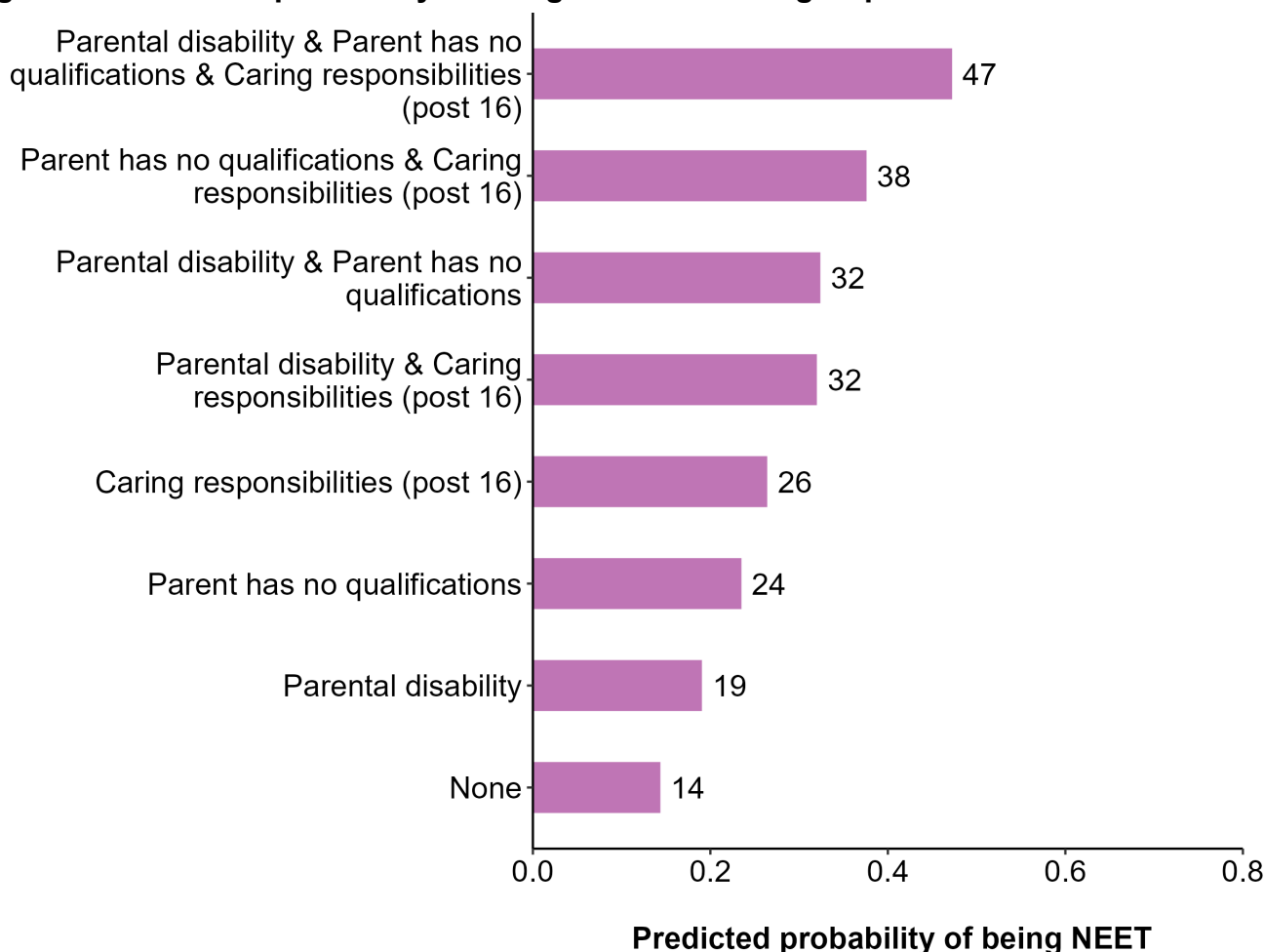
Having both SEN and achieving an academic qualification up to only Level 1 by the age of 25 was almost twice as common among men compared to women – 15% versus 8%. It was also more common among young people of white ethnicity, experienced by 13%, compared to 7% for the Mixed ethnic group – and approximately 5% for both the Asian and Black groups. It was also more common among young people who grew up in low-income households, experienced by 16% in comparison to 10% from more affluent backgrounds. Additionally, the overlap was more common in North East and East Midlands, where 17% and 18% of young people reported it, compared with London and East of England, where this figure was 7% and 8%.

Cluster 5

The fifth cluster identified a set of variables which included having a parent with a disability, neither parent having any qualifications, and the young person having caring responsibilities after age 16. The risk factors in this cluster were fairly common, with around two thirds of young people reporting at least one of the risk factors, and approximately one quarter of young people reporting at least two or more of them, but only 5% reported all three. Figure 4.8 shows how experiencing each potential combination of these risk factors increased a person's likelihood of being NEET.⁵ Each column showing the result for a different potential combination. For example, a young person might have both a parent who has a disability and has no qualifications. The figure indicates that the group who experienced all three of these risk factors had the greatest risk of being NEET, due to the sum of the different risks associated with each factor.

⁵ The predicted probability in Figure 3.4 is calculated from a logistic regression model. The predicted probabilities range from 0 to 1, with 1 meaning that NEET status is always expected to happen and 0 that it is never expected to happen.

Figure 4.8 Predicted probability of being NEET for each group in Cluster 5



Source: Next Steps Sweep 8 (age 25). Unweighted base size: 7,494.

Who is at greater risk of being in Cluster 5?

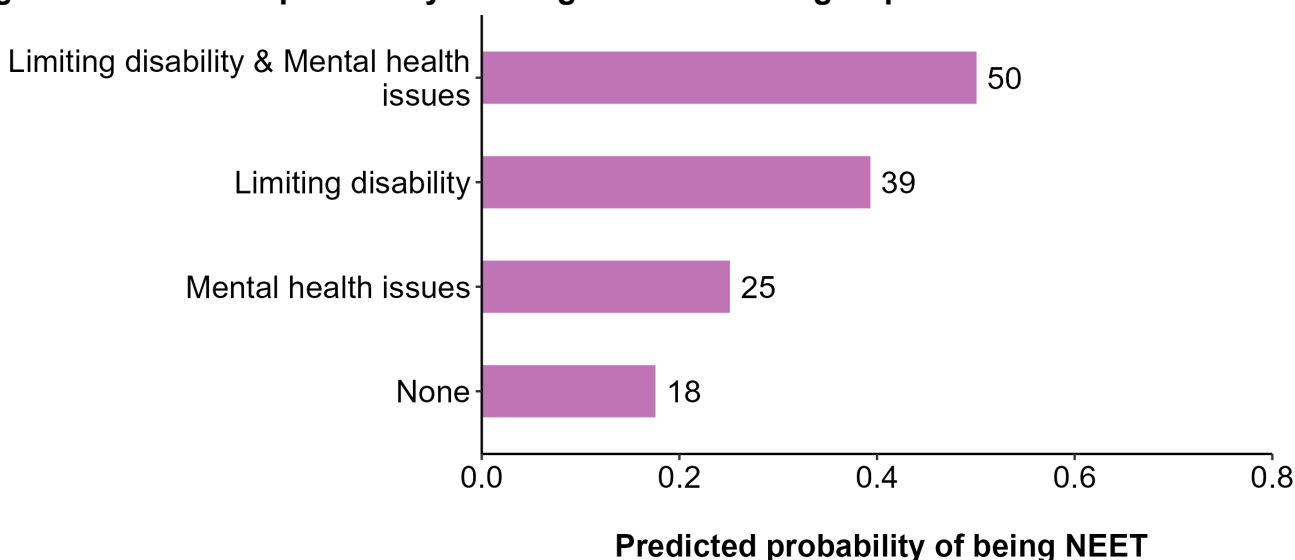
Reporting at least two risk factors from Cluster 5 was more common among young people who grew up in low-income households than among those who did not, experienced by 32% compared to 20%. It was also more common among young people of Asian ethnicity, 52% of whom experienced at least two of these risk factors compared to 27% among young people of mixed ethnicity, 22% among black ethnic groups, and 21% among the white ethnic group (see Appendix 5: Supporting tables). This is not consistent with the overall trend for the Asian ethnic group, which overall had fewer risk factors and a lower levels of NEET status than any other ethnic group. This suggests there are subgroups among the Asian population which have higher levels of risk factors, who may be of interest for policy further research and focus by policy makers. In particular, older people (65 or over) among Bangladeshis and Pakistanis are known to have poorer health, potentially responsible for the increased burden of caring responsibilities, than other Asian ethnic groups (PHE, 2021), and also to have higher levels of multi-generational living arrangements (ONS, 2023b).

Cluster 6

The final cluster was formed from the two risk factors associated with the young person's health: having had a limiting disability and having experienced depression. Reflecting the fact that poor health in one of

these areas was associated with a greater chance of having problems in the other. This overlap was experienced by around 8% of young people. Figure 4.9 shows the probability of experiencing NEET status associated with each risk factor and with their overlap. As for the other clusters, except for Cluster 6, the risk effect of experiencing these risk factors did not change when someone experienced both of them. That is to say, young people who had a limiting disability had the same increase in risk of being NEET from experiencing depression as young people without a limiting disability.

Figure 4.9 Predicted probability of being NEET for each group in Cluster 6



Source: Next Steps Sweep 8 (age 25). Unweighted base size: 7,637.

Who is at greater risk of being in Cluster 6?

The combination of experiencing both mental health issues and a limiting disability was more common among young people who identified as bisexual and gay or lesbian than among heterosexual young people: 26% of the bisexual and 15% of the gay or lesbian young people experienced both risk factors as opposed to 8% of the heterosexual young people.

Young people of white and mixed ethnicity were also more likely to experience both risk factors in comparison with black and Asian ethnic groups, experienced by 9% and 7% respectively among young people of white and mixed ethnicity, compared to 4% in the latter two groups.

5. Which risk factors were most important?

To summarise, we have seen that each individual risk factor was associated with an increased risk of being NEET – and that experiencing more than one risk factor was likely to mean an even greater likelihood of being NEET. This increased as the number of risk factors a person experienced rose. In most cases, there were not interactions between the risk factors we observed as forming clusters, meaning that the increase in risk was caused by the additive effect of the different factors, but not by a combined factor effect. That is to say, the variables which clustered together did not seem to change the importance of other variables in their cluster for how likely someone was to be NEET (the exception being SEN and not holding a qualification above Level 1). Building on this, to understand how someone's risk of being NEET was linked to each risk factor – while controlling for the effect of each of the other risk factors – a logistic regression model with all the potential risk factors was constructed. This provided an estimate of someone's likelihood of being NEET for different combinations of risk factors.

First, a separate model for each of the domains of marginalisation, including only the risk factors from that domain, were developed. This resulted in five models, one each for 1) education, 2) family circumstances, 3) health, 4) risky behaviour, and 5) living standards. These identify whether each risk factor continues to contribute significantly to the model, and how far a young person's likelihood of being NEET changes if they experience that risk factor, when controlling for only the risk factors within each domain. In these models all risk factors remained statistically significant predictors of NEET status. The one interaction effect found to be significant from the cluster variables was also included in the model for education, and this too remained significant.

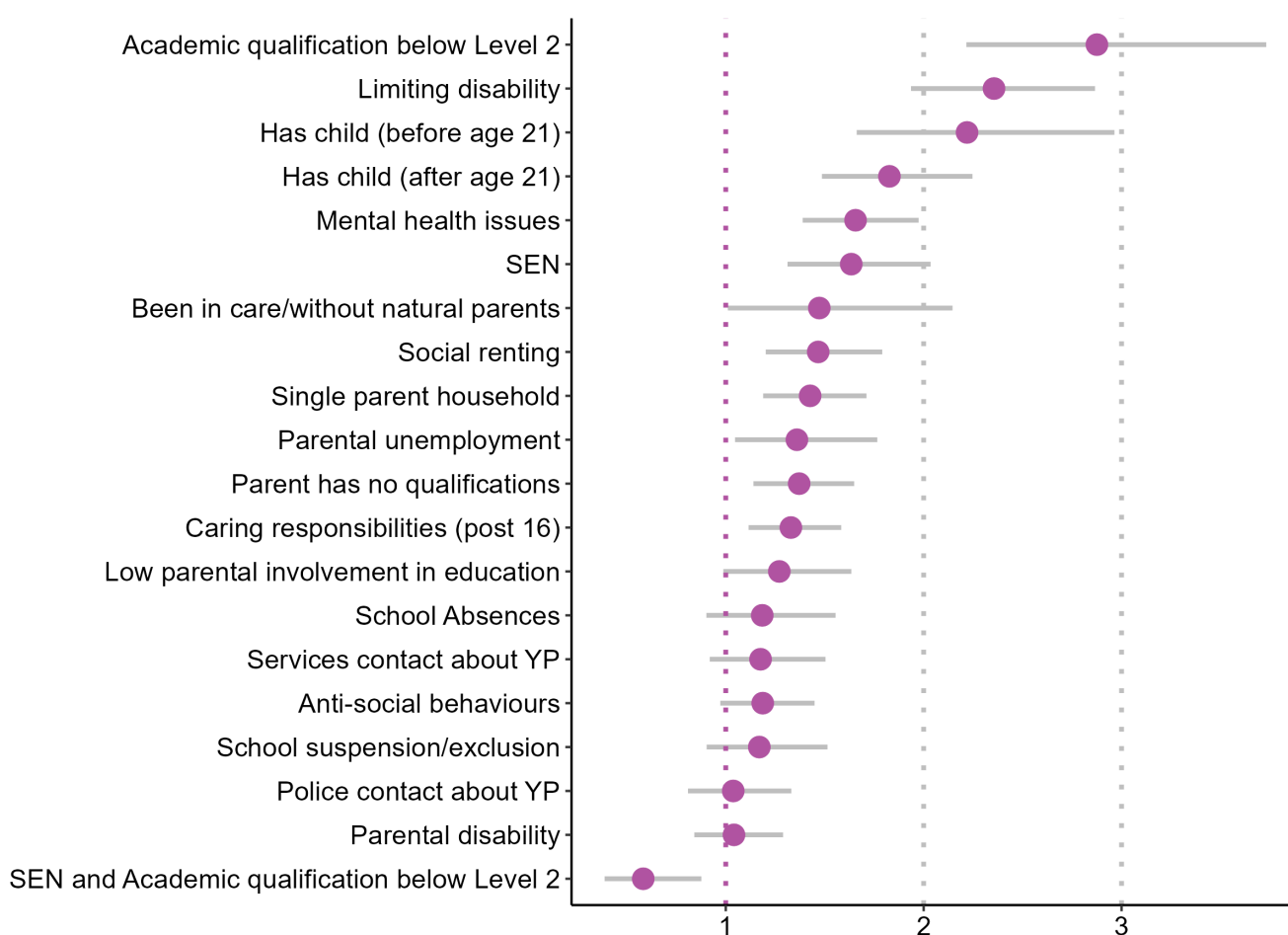
Given that all the risk factors were significant at this stage, a final combined model was then estimated that contained the risk factors from each of the domain models. In this final model several risk factors ceased to be statistically significant predictors of NEET status. This suggests that, for some of these risk factors, the differences in someone's likelihood being NEET between those that do and do not experience them is explained by one of the other variables in the model. The risk factors that ceased to be statistically significant in the final model are:

- Absence from school for one month or more in the last year,
- Expulsion or suspension from school,
- Contact with parents from educational or social services about the young person's behaviour,
- Contact with parents from police about the young person's behaviour,
- Anti-social behaviour by the young person,
- Low parental involvement in education,
- And living in a household with a disabled parent.

Aside from the final two, the underlying factor behind these variables seems to be their focus on the young person's behaviour. That these variables were consistently non-significant once other risk factors were accounted for suggests that the risk associated with behavioural variables in understanding someone's likelihood of being NEET, was largely explained by other factors in the model⁶.

The results of the final combined model are shown in Figure 5.1. The purple dots represent the odds ratios; for each risk factor, the odds ratio is an estimate of how many times greater someone's odds of being NEET were if they have experienced a particular risk factor relative to the group who has not experienced it. An odds ratio above 1 indicates an increase (for example, an odds ratio of 2 indicates that the odds of someone being NEET were two times greater if they have experienced that risk factor than if they have not), while a ratio of below 1 indicates a decrease in the odds of being NEET. These findings are controlling for all of the factors presented in the chart.

Figure 5.1 Odds ratios for being NEET



Source: Next Steps Sweep 8 (age 25). Unweighted base size: 7,091.

⁶ This might indicate a causal relationship between material deprivation and behavioural elements, but the analysis carried out in this study cannot provide statistical empirical evidence for this relationship.

Each odds ratio has a confidence interval around it, shown by the horizontal grey line. This shows the range of values that we are 95% confident will include the value for this estimate in the population of interest, based on the sample of respondents to this survey. When this line crosses the vertical dotted purple line, we cannot be confident the odds ratio is different from 1 in the population of interest, meaning that the risk associated with the risk factor might not be statistically significant.

The risk factors are ordered by the size of their odds ratio, showing that relative to those with an academic qualification at Level 2 or above, those without such a qualification had the largest increase in their odds of being NEET, with the odds approximately three times as great.

Overall, Figure 5.1 suggests that, when controlling for the other risk factors in the model, the most important determinants of risk for NEET are:

- Not having an academic qualification above level 1.
- Having a limiting disability.
- Having own child before the age of 21.
- Having own child between ages 21 and 25.
- Having a poor mental health condition.
- Having been identified as SEN.

These are all factors found to be likely to overlap with each other in the cluster analysis, as they all belong to either Cluster 2, Cluster 3 or Cluster 5, meaning that young people were relatively likely to experience these influential risk factors in pairs, and to suffer from their additive risk.

6. Developing a risk index for being NEET

The final stage of this analysis was to identify if some groups in the population were more at risk of being NEET than others, based on the final regression models findings about which risk factors were the most important predictors of someone's likelihood of being NEET. This was achieved by creating a risk index, summarising a person's risk of being NEET based on their unique combination of risk factors, and profiling the risk index by a range of different groups in the population to see if the mean risk index score differed for any of these groups. The groups explored included a number of characteristics protected by the 2010 Equality Act, but also a number of other factors of policy interest:

- **Country of birth:** whether the young person was born in the UK or elsewhere.
- **Sex:** whether male or female.
- **Sexuality:** whether young person thinks of themselves as heterosexual, gay, bisexual or 'other'.
- **Government Office Region:** the region of England where the interview was conducted at Sweep 8.
- **Household income:** whether the household the young person was living in fell into the bottom 10% of households ranked by income from Sweeps 1 – 4.
- **Ethnicity:** classification of young person's ethnicity into 15 detailed groups and 5 more general groups.

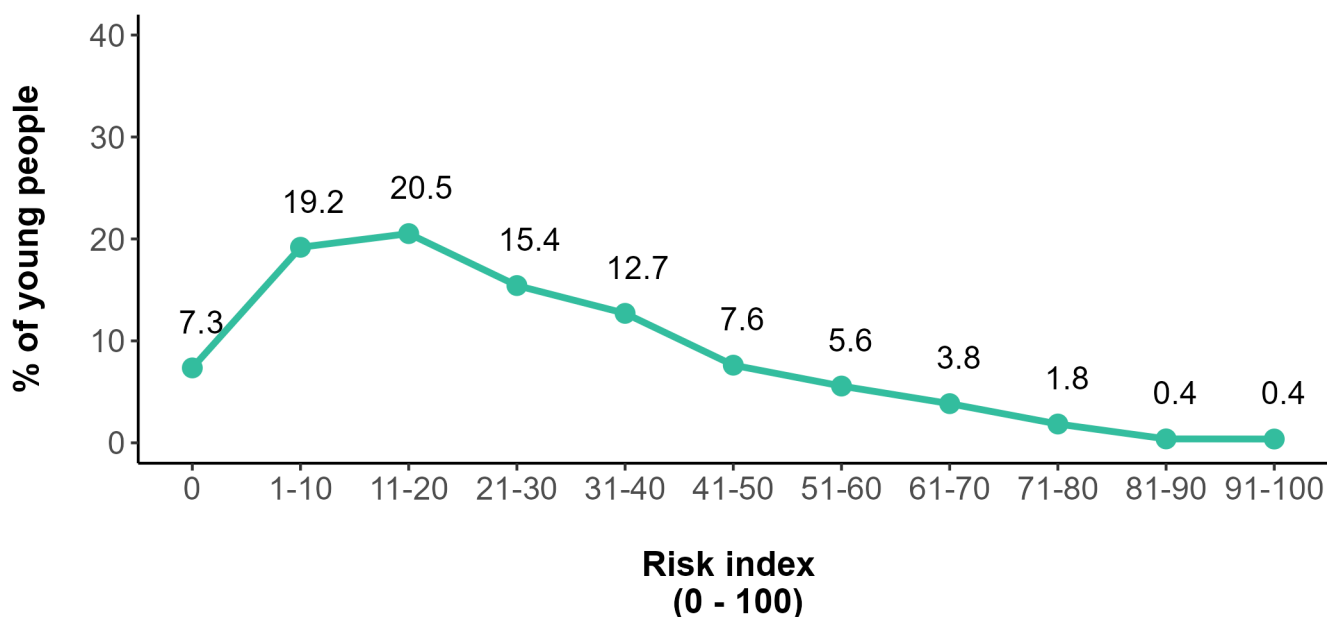
Creating an index of risk

The risk index for young people's likelihood of being NEET is a risk score assigned to each person in the Next Steps dataset which ranges from 0 – 100, indicating their relative risk of having been NEET at least once from age 18 to 25, with 100 showing the highest level of risk. It was created summing all the risk factors each study participant experienced; the sum was weighted by the factors' relative risk importance, as identified in the final model presented in Chapter 5. The results were projected on a 0-100 scale to ease interpretability⁷.

The NEET Risk Index is therefore a measure of the relative risk that each person has of being NEET, given the 19 factors included in the analysis. As is shown in Figure 6.1, most young people were clustered around the lower end of the NEET Risk Index. Nearly half (47%) were assigned a risk score lower than 20, and another 36% a risk score of up to 50. Only 12% of young people were assigned a risk score greater than 50, with the highest scores above 80 reported by less than 1% of young people.

⁷ More details on the production of the NEET Risk Index can be found in Appendix 2.

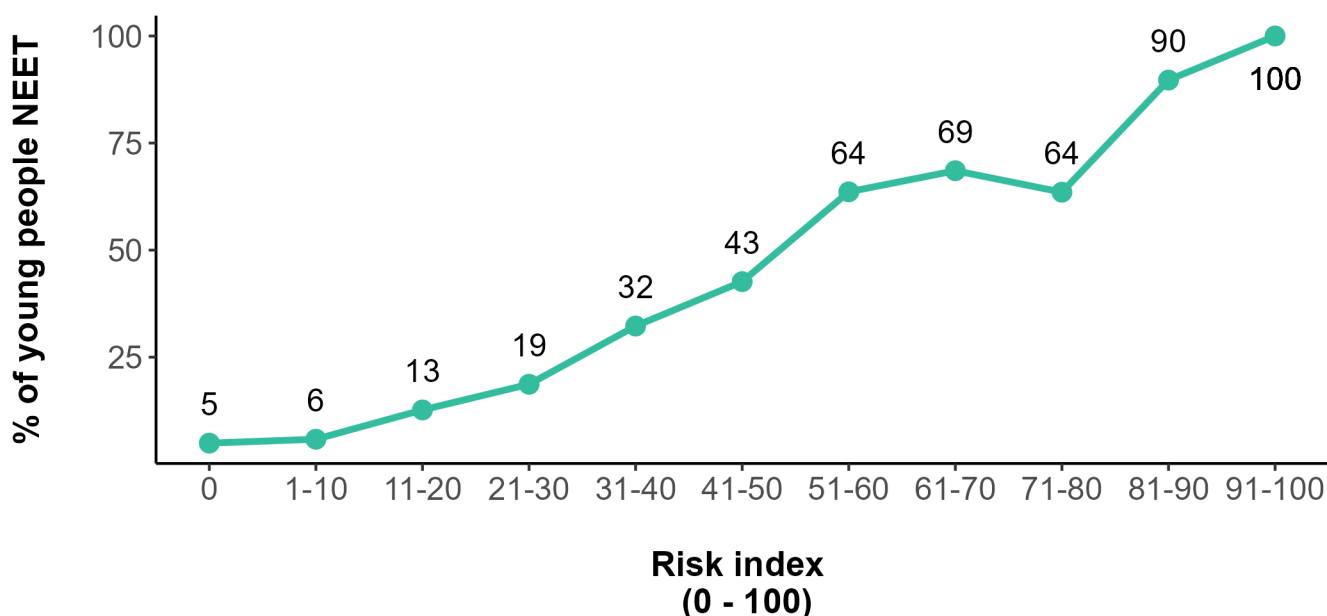
Figure 6.1 Percentage of young people assigned to each level of the NEET Risk Index



Source: Next Steps Sweep 8. Unweighted base size: 7,091.

Figure 6.2 is included as an illustration that the NEET Risk Index is providing an effective measure of the likelihood of young people experiencing NEET status. When comparing the proportion of young people who have been NEET with their risk score, we can see that the proportion of young people who were NEET increases from 5% in the group with a score of zero on the risk index to 90% in the group with scores of between 81 and 90 and 100% in the group with a score of 91 or higher.

Figure 6.2 Percentage of young people who had been NEET by Risk Index score



Source: Next Steps Sweep 8 (age 25). Unweighted base size: 7,082.

However, the NEET Risk Index is not a perfect measure of risk: it was built around 19 risk factors, but there are other important characteristics which affect how likely they were to be NEET which were not considered in this analysis. There is also a small drop in the percentage of young people who have been

NEET in the group with a score between 71 and 80. This reflects the fact that at the upper end of the risk index there were only a few young people in our sample with a score this high, which means the estimate of how many of them are NEET is less precise.

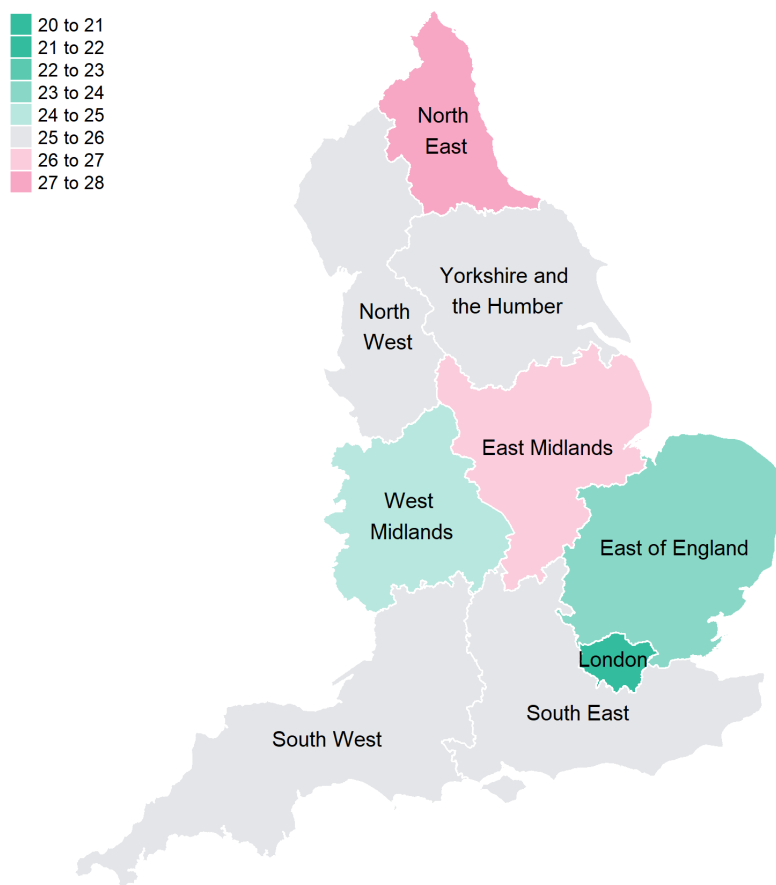
Which groups of people had the highest NEET Risk Index score?

- **Country of birth:** among young people born in the UK there was a slightly higher mean NEET Risk Index score than among those born outside the UK, 25 compared to 23 points. However, this difference was not statistically significant meaning we cannot be confident that this difference did not arise due to sampling variability.
- **Ethnicity:** comparing first five general categories of ethnicity, white, black, Asian, mixed ethnicity and 'other', young people of Asian ethnicity had a lower mean score (of 19) than the other groups. Young people of black ethnicity had the highest score at 27, while the mixed and white ethnic groups both had a mean score of 25. Those in the 'other' ethnic group are not reported, due to their small base size.
 - Within the Asian ethnic group, however, young people of Pakistani and Bangladeshi origin had a higher mean score of 23, higher than those of Indian ethnicity and those from 'other' Asian ethnic groups, with mean scores of 15 and 17 respectively.
- **Household income:** the mean NEET Risk Index score was 32 among young people who grew up in low-income households, while for young people who grew up in more affluent households the mean score was substantially lower at only 22.
- **Sexuality:** young people of heterosexual and gay or lesbian sexual orientation had lower scores on the NEET Risk Index, of 24 in both these groups compared to 33 for young people who identified as bisexual or 'Other' when asked about their sexuality.
- **Gender:** the mean NEET Risk Index score was slightly higher for female young people than for male, 24 compared to 26 points. Although small, this difference was statistically significant.

Government Office Region

Figure 6.3 shows regional trends in the NEET Risk Index. Regions are coloured so those with a high mean risk index score are a darker pink, while a darker green indicates lower values on the score and therefore a lower risk of being NEET. The two areas of highest risk are all areas of relatively high deprivation in England, while the mean index score was lowest for London, reflecting the relatively high levels of economic opportunity there. However, this analysis is based on large geographical areas which will have considerable variation within them. All will include pockets of deep poverty and economic disadvantage, as well as areas of affluence where young people have a low risk of being NEET. A more detailed exploration of the relationship between geographical regions and risks of being NEET among young people will show a more nuanced picture of how geography impacts young people's economic activity and education.

Figure 6.3 Mean risk index score by region



Source: Next Steps Sweep 8 (age 25). Unweighted base size: 7,063.

7. Discussion

This research project approached marginalisation as “a state of disadvantage, exclusion, or transgression from a ‘norm’ or typical situation/ social model, as defined by wider society” (Addario et al., 2023). It explored it as a form of multidimensional disadvantage which could be present in multiple areas of someone’s life, with the potential to impact the employment and educational opportunities available to people. This framework was supported by the empirical evidence explored in this study. The different dimensions of disadvantage, described in this report as risk factors, were all associated with an increased likelihood of experiencing NEET status. It also emerged clearly that risk factors do overlap. Only 8% of the young people had zero risk factors, and 14% experienced one. This means that 75% experienced overlapping dimensions of disadvantage, and there was an associated increase in the risk of facing NEET status from having a combination of risk factors.

In addition, by exploring the links between risk factors in combination, once the role of others is taken into account, our study provides some evidence to help policy makers and frontline staff understand which dimensions of marginalisation are of highest priority for action when working to reduce the level of NEET status in young people. Specifically, we identified six risk factors that strongly associated with NEET status:

- Not having an academic qualification above level 1.
- Having a limiting disability.
- Having own child before the age of 21.
- Having own child after age 21 and before 25.
- Having a mental health condition.
- Having been identified as SEN.

These highlight the importance of public initiatives which support people with these characteristics, such as education, programmes to promote mental wellbeing, and income support for people with disabilities, SEN and childcare responsibilities. It should be noted, however, that the latest timepoint in this data, when these young people were 25, was collected in 2015. The ages when young people’s NEET status was analysed ran from 18 to 25 and were collected between 2008 and 2015. This means that some of the difficulties which young people face may have changed to some extent since then.

We also found some evidence of clustering between different types of risk factor, with some risk factors relatively more likely to occur alongside specific other types of marginalisation. Our analysis showed the presence of six different groups of risk factors, but also that these groups were quite weak – meaning that the borders of the clusters were blurred. This represents a potential challenge for policy initiatives; the heterogeneity of the type of risk factors young people experience make it difficult to target policy effectively. That is to say, the combination of risk factors young people experience varies greatly, making

it difficult to apply a uniform set of policies effectively, as they will need to tackle several dimensions of disadvantage and the specific set of characteristics a young person will have is not consistent.

Nevertheless, our analysis identified at least two clusters that could be considered a potential focus, given they generate a substantial increase in risk due to the presence of four of the six most influential risk factors, and that these overlaps of risk factors were relatively common in the population:

- Having been identified as SEN and not having an academic qualification above level 1 (experienced by 12% of the population of young people), and
- Having a limiting disability and poor mental health (experienced by 8% of young people).

Our research has also shown that some groups of young people were at greater risk of being NEET, due to the different types and numbers of risk factors they faced. This implies that multidimensional disadvantage, or the risk of being severed from mainstream social processes, is unequally distributed in society. Particularly at risk were those that grew up in low-income households, young people of black or mixed ethnicity, as well as people of Pakistani and Bangladeshi origin, and people who identified their sexuality as homosexual, bisexual or 'other'. However, although some groups in the population might be systematically more likely to be in a position of relative disadvantage compared to others, our analysis also shows that within all groups the distribution of risk is unequal. That is to say that even among groups of young people, such as those of Indian ethnicity, who were relatively unlikely to have multiple risk factors, there were still many people who had a high NEET Risk Index score. Given this, when seeking to target policy it may be more effective to direct intervention towards people with specific risk factors, rather than target everyone in a particular group, such as those who grew up in low-income households, as not everyone in this group will experience a high risk of being NEET.

Limitations and further research

Our study has also some important limitations and should be considered as a steppingstone for further analysis.

Context

The first limitation is the context of our study: the young people in this sample turned 18 at the time of the 2008 financial crash, which started a period of economic downturn and was followed by a decade of austerity in public spending. This means the risk of experiencing NEET in our cohort could have been influenced by these events, and some of these risk factors may be specific to that context. The prevalence and importance of the risk factors identified in this study may also have changed since this time period, and new risk factors have become important which are not measured here.

This issue of context could be addressed by repeating the current analysis with a more recent dataset. The successor to Next Steps, called "Our Future", began data collection in 2012/13 with a new cohort of 13- to 14-year-olds. It includes many of the same measures and could be used to explore whether the risk factors identified in this report have continued to affect young people, whether they have done so to a comparable degree, and whether new factors have arisen since then. However, key data in this study

to understand NEET status were collected during the COVID-19 pandemic when normal labour market conditions were disrupted, which would limit comparability between studies.

Indicators of marginalisation

The second limitation is the type of risk factors that could be identified in the Next Steps data and the exclusion of those factors that could not be effectively included in this analysis. We were not able to explore a number of factors that were found to be particularly important in the existing literature, such as experiences of crimes in a young person's local area, living in an area with high levels of deprivation, and parental attitudes towards education.

A first step to address this limitation would be to extend the analysis to make use of the linked National Pupil Database information which can be accessed on Next Steps participants. This would provide detailed records of these young people's academic attainment at different stages in their education. A related possibility would be to make use of developing administrative datasets which follow young people over time through their education and into employment (in particular, the Longitudinal Educational Outcomes study would be suitable). These would provide a detailed record of young people's employment outcomes and education, as well as some other risk factors such as having special educational needs. It would not, however, include things which are only measured through surveys, such as parental attitude to education. This reflects a general difficulty in analysing the overlaps between different risk factors – that no single dataset covers all characteristics of potential interest. This means that while attempting to replicate or extend this analysis through other sources would add valuable detail, they will all have limitations specific to them.

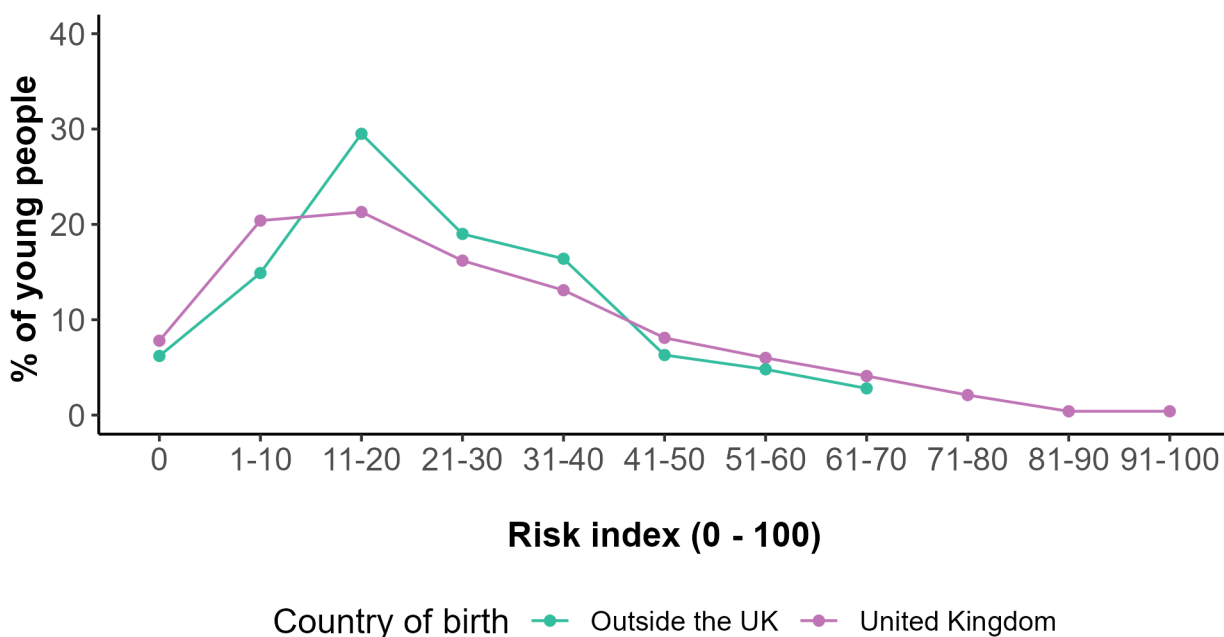
The role of time

A third limitation is the treatment of time in our analysis. The outcome variable, having had an experience of NEET at either Sweep 6, 7 or 8, did not take account of when the experience of NEET status had happened, or whether it was of a temporary or persistent nature. Similarly, we did not focus on the specific time at which a risk factor (the dimensions of disadvantage explored in our study) was experienced by study participants, but rather on their presence or absence in the course of the young adult's life. A greater focus on the timing of events can help target policy intervention to a specific period of the young adult's life, when it can be the most effective.

8. Appendix 1: Additional NEET Risk Index Profiles

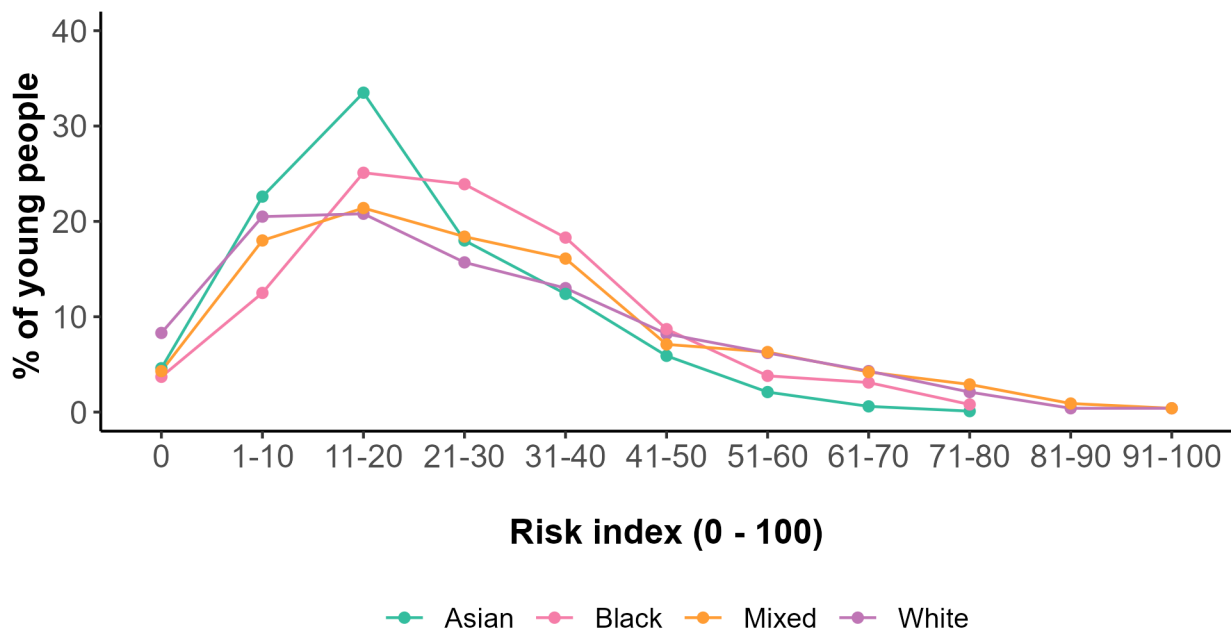
Appendix 1 presents in more detail some the findings around protected characteristics and the NEET Risk Index. While the main report focussed on the NEET Risk Index's mean score, the charts in this appendix show the distribution of the score for different groups. For example, Figure 8.1 shows the percentage of young people at each level of the index, split by whether they were born in the UK, or not.

Figure 8.1 NEET Risk Index score by country of birth



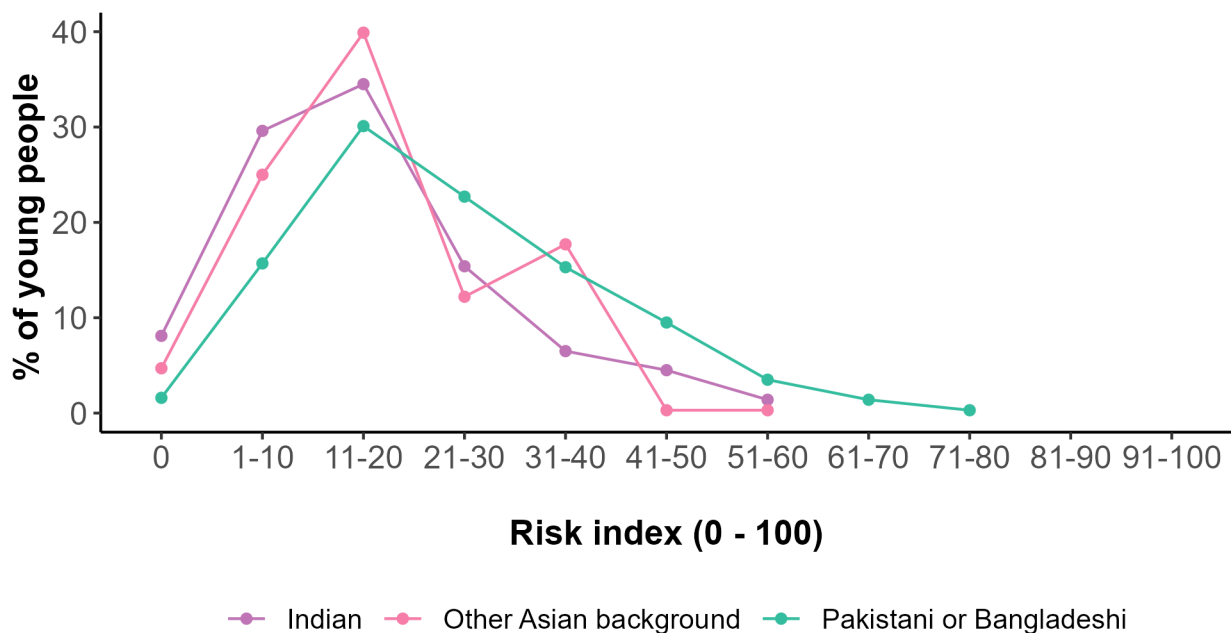
Source: Next Steps Sweep 8. Unweighted base size: 6,952.

Figure 8.2 NEET Risk Index scores by ethnicity



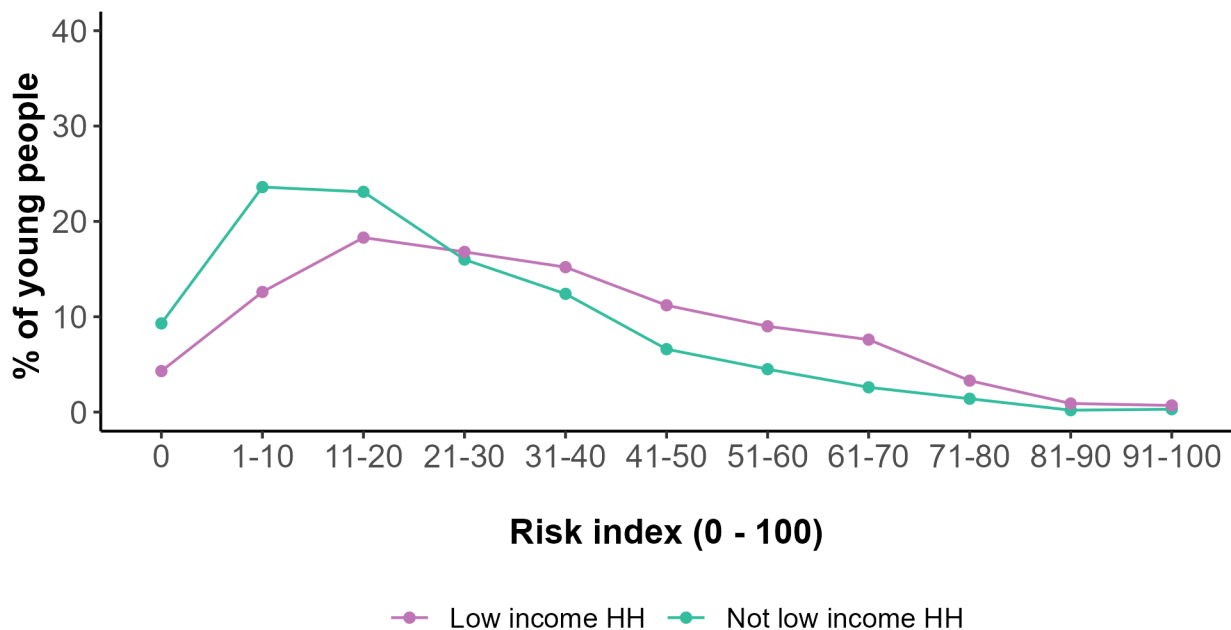
Source: Next Steps Sweep 8. Unweighted base size: 7,042.

Figure 8.3 NEET Risk Index scores by subgroups within the Asian population



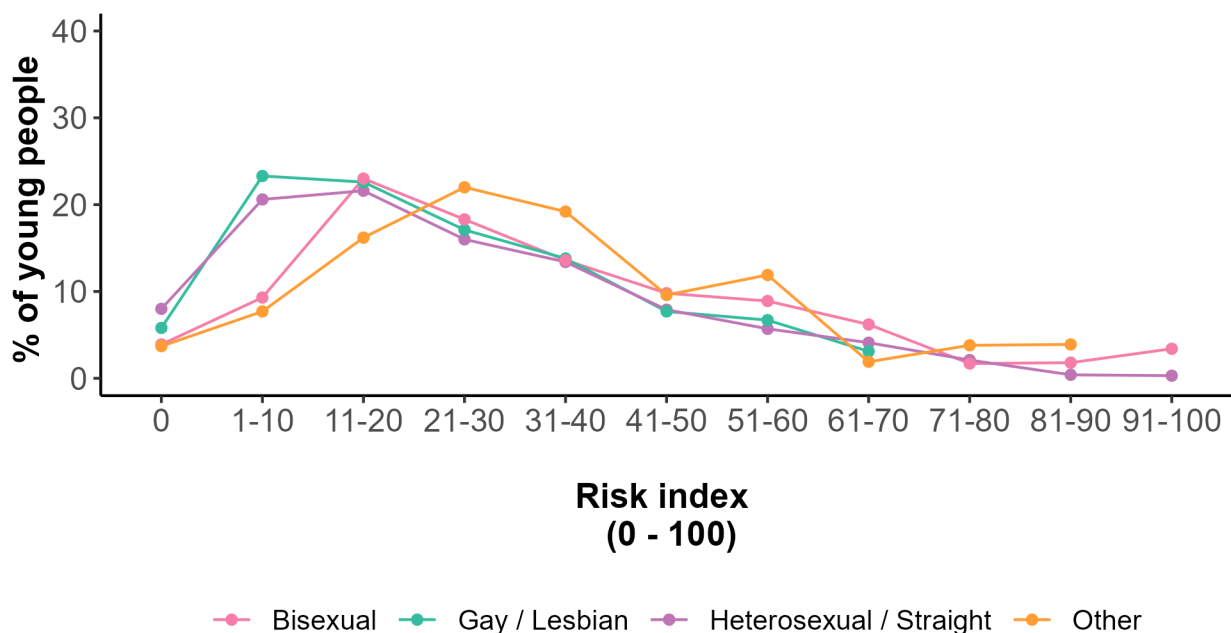
Source: Next Steps Sweeps 1 – 8. Unweighted base size: 1,014.

Figure 8.4 NEET Risk Index scores by HH income in childhood



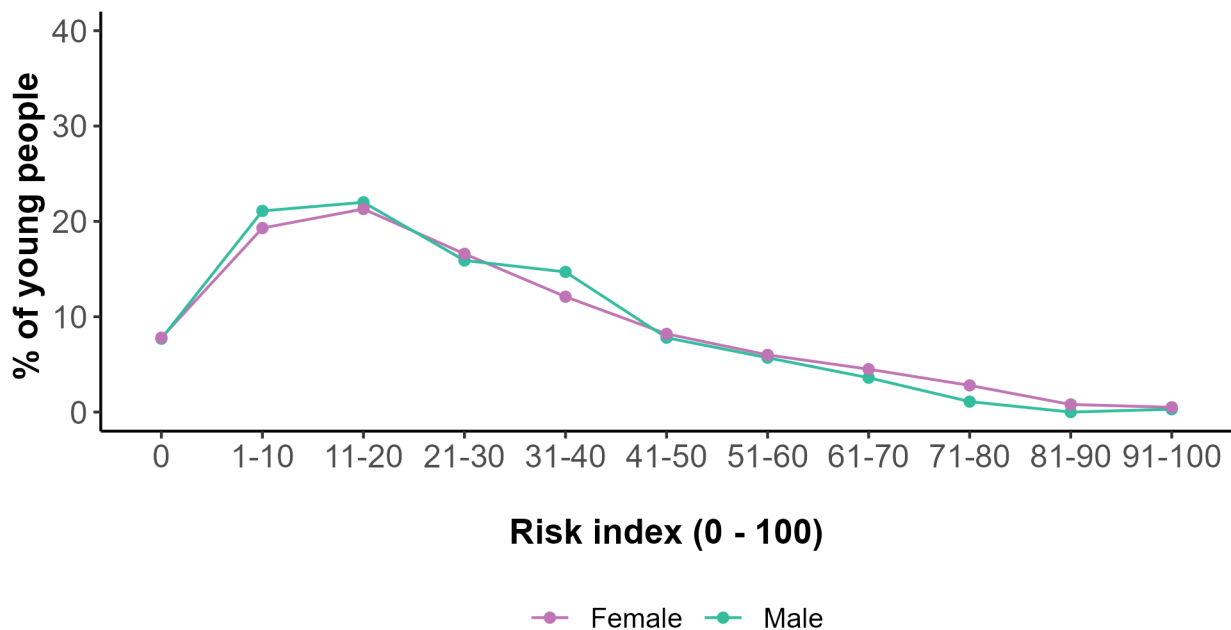
Source: Next Steps Sweep 8. Unweighted base size: 6,972.

Figure 8.5 Proportion of young people with different NEET Risk Index scores by sexuality



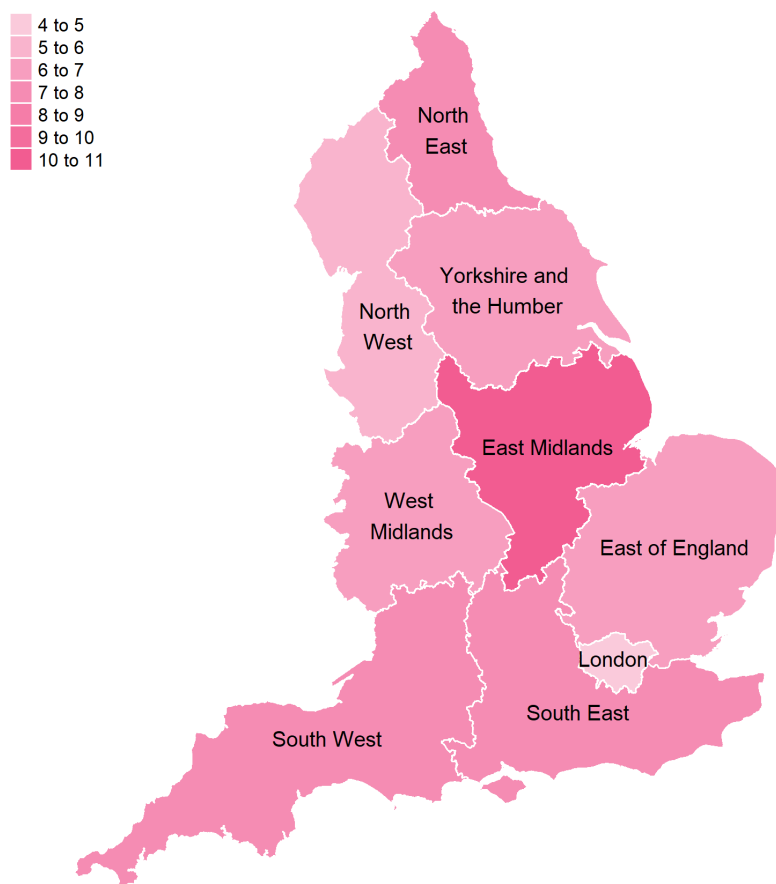
Source: Next Steps Sweep 8. Unweighted base size: 6,881.

Figure 8.6 Proportion of young people with different NEET Risk Index scores by sex



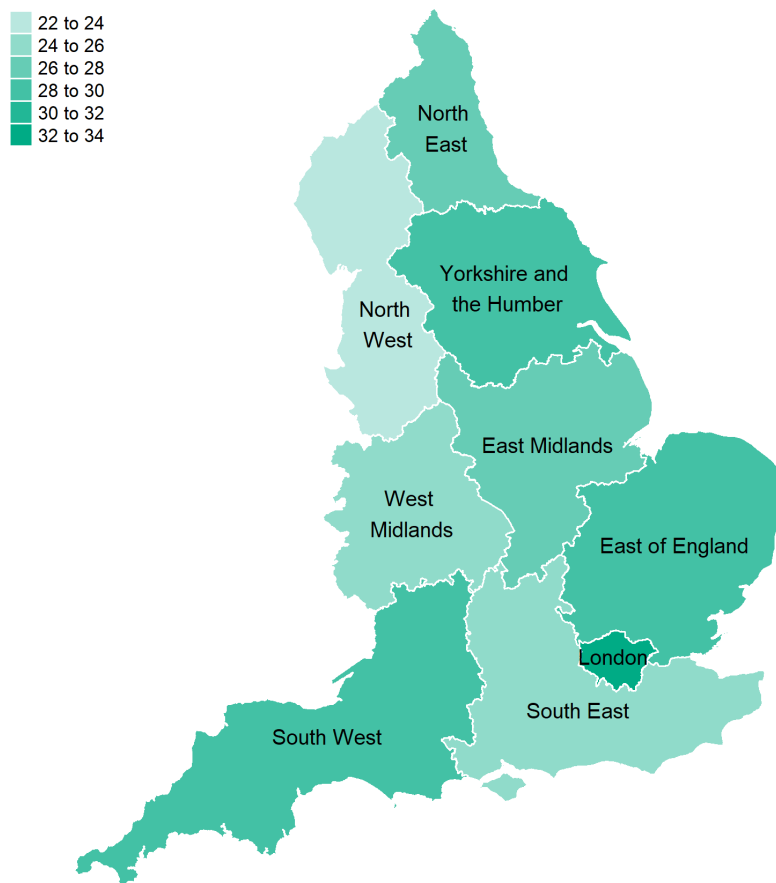
Source: Next Steps Sweep 8. Unweighted base size: 7,091.

Figure 8.7 Percentage of young people with risk index score above 60 by region



Source: Next Steps Sweep 8. Unweighted base size: 7,063.

Figure 8.8 Percentage of young people with risk index score below 10 by region



Source: Next Steps Sweep 8. Unweighted base size: 7,063.

9. Appendix 2: Measurement of marginalisation

The creation of the risk factor variables combined data from different questions and multiple Next Steps sweeps. Table 9.1 below shows the timepoints at which each of our indicators was collected, with most falling in the first 4 timepoints – between ages 13/14 and 16/17. This means that most of our risk factors were collected before our outcome of interest, having been NEET at least once, which is measured in Sweeps 6, 7 and 8. These sweeps are highlighted in purple to highlight which risk factors were collected at the same timepoints as the outcome.

Where a timepoint is indicated by a 'B' instead of a tick, data was used from the ethnic minority boost sample. This was a small additional sample (approx. 130) of young people from ethnic minority backgrounds introduced at Sweep 4 to increase the study's ability to report reliable findings for ethnic minorities. These young people were often asked the questions we have used at a different age to the rest of the sample, which may affect how they answered them. For example, the boost sample was not asked about being identified as having special educational needs at Sweeps 1 and 2, they were asked at Sweep 4. This means there was more time for them to be identified as SEN, and so the prevalence among this group may be higher than it would be if they had been asked at the same age as other respondents.

The specific approach to deriving each indicator of marginalisation is described in more detail below, however, in general we created simple binary measures that indicated whether someone had experienced a risk factor **at least once** across the timepoints it was asked at. For example, data on disability status and long term-health conditions were collected at three time points and if a child reported them at any of these then they were coded as having this risk factor.

This allowed for a relatively simple operationalisation of risk factors which could look across multiple types of marginalisation, however, it does introduce certain limitations to the results. It removes some of the detail that could be gained from exploring whether, for example, someone's age when they experience a risk factor matters. Potentially, the role of age could also vary between risk factors. Similarly, it does not explore whether the duration of a marginalising experience matters, for example, whether someone consistently had poor mental health over several years while a child, or experienced poor mental health for a shorter period of time.

The most important limitation for interpreting these results is that by using data from different Sweeps of the Next Steps study, there is a potential for the underestimation of how common risk factors were. This is because at each Sweep we have some people who did not respond to either the entire survey or who did not respond to the particular questions we are interested in. For example:

- Questions about having a disability or a limiting health condition were asked in three Sweeps of Next Steps.
- Sample members were considered in our analysis to have a disability or limiting health condition as long as their responses was 'yes' to the disability question in at least one Sweep of the study.

- Sample members who responded in at least one Sweep of Next Step and their answers in the Sweeps where they took part were consistently 'no' were coded as not having any disability of limiting health conditions.

Some of the sample members considered not to have a disability of a limiting condition may have been coded incorrectly. This is because in one of the Sweeps where they did not answer, they may have actually been a 'yes' if they had responded to the survey. This means that we may be underestimating the total number of people who have risk factors when we operationalise them in this way.

Although this introduces some level of underestimation, we have decided to adopt this approach because it maximises the use of available information. Across the 19 risk factors we have identified, this approach reduces the loss of cases to about 8% of the sample available for analysis (616 cases have at least one missing value, out of 7,707 cases), the variable with the highest level of missingness has 6.2% of missing values (479 cases), and only 5 variables have missingness above 1%. By contrast, if we only used cases with complete data, 4,725 cases (61%) would be lost from the analysis. This would mean a loss in sample size and statistical power, as well as introducing bias into the analysis by excluding cases less likely to respond across all waves and to all items with a wave. These people may well have systematic differences to those who are more likely to respond consistently to the survey and excluding them from the analysis would potentially bias the results.

Table 9.1 Risk factors for NEET status

Risk factors	Sweep number and young person's age							
	1 13/14	2 14/15	3 15/16	4 16/17	5 17/18	6 18/19	7 19/20	8 25
Education								
SEN	✓	✓		B				
Absence from school	✓	✓	✓					
Expelled / suspended from school	✓	✓	✓	B				
Parental involvement in education	✓	✓	✓					
Highest attained qualification								✓
Family circumstances								
Caring responsibilities post-16				✓	✓	✓	✓	
No parental qualifications	✓	✓		✓				
Parental disability	✓							
Single parent household	✓	✓	✓	✓				
Had a child (pre-21)								✓
Had a child (ages 21-25)								✓

Lived in care / apart from birth parents	✓	✓	✓	✓				
Health								
Disability status	✓	✓		✓				
Mental health		✓		✓				✓
Living standards								
Social housing	✓	✓	✓	✓	✓	✓	✓	
Parental unemployment	✓	✓	✓	✓	✓			
Justice and personal safety								
Risk behaviour	✓	✓	✓	B				
Contact from social or educational welfare services	✓	✓	✓	B				
Contact from the police	✓	✓	✓	B				

Operationalising marginalisation

This section describes the operationalisation of each of the indicators of marginalisation used in this analysis, grouped by their domain – education, family life, living standards, education, and risky behaviour. Wherever we refer to the ‘main parent’ this is the parent who was reported to be most involved in the child’s education.

Education

- Special educational needs (SEN): this indicator shows whether a “young person has ever been identified, either by their school or by someone else, as having special educational needs of any kind” at either Sweep 1 or 2 (or at Sweep 4 in the ethnic minority boost sample).
- Absence from school: this indicator shows whether a young person has been absent from school for a “continuous period” of one month or more in the 12 months preceding the survey, collected in Sweep 1 to 3. This excludes school holidays.
- Exclusion or suspension from school: this indicator shows whether a young person has been temporarily suspended or excluded, permanently excluded or expelled from school, since the start of school year. It was collected in Sweeps 1 to 3 and in Sweep 4 for the ethnic minority boost sample.

There are two source questions to this variable: “Has [the young person] ever been permanently excluded, that is expelled from school for good?” and “Has [the young person] ever been temporarily excluded, that is suspended, from a school for a time”. Young people are coded as positive at this variable if the response to either of these was ‘yes’. The vast majority of these cases are suspensions, which means the variable will primarily reflect the impact of suspension, not the impact of expulsion. For example, in Sweep 1 fewer than thirty people were expelled, while around 500 were suspended.

It should also be noted that for the ethnic minority boost sample, collected in Sweep 4 (n = 132) this variable only measures exclusion/expulsion in that year, whereas for the rest of the sample it measures exclusion or suspension over a period of 3 years and at a younger age.

- Highest qualification below Level 2: this measures whether highest academic qualification achieved by the young person at age 25 is below NVQ Level 2. This is a derived variable based on a combination of survey responses and National Pupil Database linked qualification data for Key Stage 4 and Key Stage 5.
- Parental involvement in education: whether the main parent feels involved in their child's school life. It compares those who choose "very involved", "fairly involved" or "not very involved" to those who say "not at all involved" (Sweeps 1 – 3). It is also collected for the ethnic boost sample at Sweep 4.

Family situation

- Caring responsibilities (post-16): whether the young person has caring responsibilities, either within or outside the household after the age of 16 (collected in Sweeps 4 – 7). This includes the responsibility to "regularly look after any ill, disabled or elderly relatives or friends aged 15 or more and in need of care, without being paid" and to "regularly take care of any children aged 14 or under".

There was a quite high level of missing data on this indicator (6,626 valid cases out of 7,707). To address this, for those cases where there was missing data a proxy was used. If respondents had either a disabled relative, a child or they had caring responsibilities before the age of 16, they were coded as having caring responsibilities post-16. Those who had none of these were coded as having no post-16 caring responsibilities.

Having a disabled relative and having a child are other indicators described in this section, however, having caring responsibilities before age 16 was derived specifically for this indicator. It measures whether a young person has caring responsibilities within the household (Sweeps 1-3). This is care on a "regular basis" for family or someone else, delivered in the household.

- Parental education: whether **either** parent held any qualifications when the young person is age 15 (Sweep 3). If there is missing data at this sweep, then their highest qualification is taken from an earlier sweep (Sweeps 1 or 2).
- Parental disability: whether **either** parent has a longstanding illness, disability, or infirmity which limits their activity (Sweep 1).
- Having been in care or lived apart from their birth parents: this can be considered a measure of disruption to the young person's family, and had two components:
 - The care element of this indicates whether a young person either lived with foster parents, in a young people's home, in local authority care, or has been placed for adoption from Sweeps 1 and in Sweep 4 for the ethnic minority boost sample. This means that for most of our sample it measures having been in care *before* Sweep 1 (age 13/14).
 - Whether a young person has spent time apart from their natural parents is based on whether the young person was **not** living with their natural parents at the time of the interview in either Sweep 1, 3 or 4.
 - The indicator for being in care had a high number of missing values (around 700). To reduce this, when these were combined, as long as respondents had valid data for one of these measures they

were coded as present. Only respondents with missing data at both indicators were coded as missing.

- Single parent household: this is based on a derived variable in the Next Steps dataset which indicates whether the young person was living in a single parent household in Sweeps 1, 2, 3 or 4.
- Whether young person has own children: in Sweep 8 the young person was asked if they have any children and, if so, what their age was. This was used to create two variables which indicate whether someone has a child aged 0 – 4 (born after between age 21 and 25) or aged 5 – 11 (before age 21).

Risky behaviours

- Young person's behaviour: whether the main parent has been contacted by social services / educational welfare services / any other similar services about their behaviour (collected at Sweeps 1, 2, and 3).
- Contact with police: whether police have been in touch with the main parent about something the young person has done (collected at Sweeps 1, 2, 3 and 4).
- Criminal or anti-social behaviour: whether the young person has ever graffitied on walls or vandalised public property or shop-lifted or been involved in a disturbance (e.g., at a football ground, a railway station, music festival, riot) (collected at Sweeps 1, 2, 3, and 4).

Living standards

- Social housing indicator: whether the young person has ever lived in social housing / council housing in any Sweep (1 – 8).
- Parental unemployment: whether **either** parent had been unemployed at the time of the survey from Sweep 1 to 4. This includes parents whose main or secondary current activity was being “unemployed or/and looking for a job”. This measure was also collected at Sweep 5, however, parental unemployment at this wave is unusually high. There are approximately 400 cases where both parents are unemployed, and 1,900 where at least one parent was. This Sweep was conducted in 2008, at the time of the financial crisis, so this likely reflects a real increase in unemployment. However, it introduces a separate issue that unemployment under the circumstances of the financial crisis may be a different type of predictor to that available in previous sweeps. It reflects external economic circumstances rather than measuring the economic marginality of the parents. Given this, Sweep 5 was excluded from the analysis.

Health

- Experience of depression: whether the young person was depressed at the time of the interview, derived from the General Health Questionnaire (collected at Sweep 2, 4 and 8). This was calculated using the GHQ-12 score. This ranges from 0 – 12 and a score of 4 or higher is used to identify possible psychiatric disorders (James et al., 2013). The questions are phrased to ask about a person's functioning in the past few weeks, so disorders identified by it may be of varying duration, and it is not

intended to identify permanent or enduring conditions. Given this, we take data from the latest sweep of the questionnaire, (Sweep 8 - age 25), and only used the earlier sweeps for this measure where data was missing at age 25.

- Disability status: whether the young person has ever had a longstanding illness, disability, or infirmity which limits their activity (collected in Sweep 1, 2 and 4, and again in Sweep 8).

The question wording changed slightly in this indicator from Sweeps 1, 2 and 4, where it asked about a disability or condition that limited people's ability to their schoolwork, while in Sweep 8 it asks whether it impacts on their ability to carry out their day-to-day activities.

10. Appendix 3: Defining NEET status

This study follows the definition of NEET status set out by the Office for National Statistics (ONS, 2016). A person is considered NEET if they do **not** meet any the following criteria:

- **Employment**

- They are currently in employment. Employment is defined as the individual being in some form of paid work either part-time or full-time.

- **Education and Training**

- They are enrolled on an education course, apprenticeship, government supported employment or training programme, or on job related training.

or

- They are enrolled on an education or training course outlined above but are waiting for the (re)start of a term or course.

To measure this definition of NEET status in the Next Steps data slightly different approaches had to be taken in each Sweep due to differences in the survey questions asked at each interview. A very similar question was asked to identify the young person’s current main activity; however, different answer options were available at each Sweep.

The table below indicates how these were coded into either NEET or not NEET.

Table 10.1 Measuring NEET status in young people

	NEET	Not NEET
	<i>All activities were in this group:</i>	<i>At least one activity was:</i>
Sweep 6	Looking after the family and home, unemployed and looking for work, waiting for exam results or result of job application, doing voluntary work.	Doing a course at university, in education, in paid work, on a training scheme or course, doing an apprenticeship, waiting for a course or job to start, or spending part of the week with an employer and part of the week at college.
Sweep 7	Looking after home/family, unemployed and looking for work, voluntary work, travelling, break from college or work, and ill or disabled	University, school/college education, paid work, training course/scheme, apprenticeship, waiting for a course or job to start, part time job and part time college, and government employment programme.

Sweep 8	In unpaid or voluntary work, unemployed, sick / disabled, looking after the home or family, maternity leave, travelling, and 'other' activities.	Employed, self-employed, working but employment status is unclear, in education (school, college, or university), on an apprenticeship, on a government training scheme.
---------	--	--

To ensure NEET status was derived more precisely, data for each type of activity is available for Sweeps 6 and 7. The identification of NEET status in Sweep 8 is based on 'main' activity and "secondary activities" only.

11. Appendix 4: Analysis methods

In the main report we restrict the analysis methods to what is necessary to interpret and understand the results. In this appendix we will explain them in more detail, in particular the cluster analysis, the regression analysis, and the creation of the risk index.

Cluster analysis

Cluster analysis is a descriptive method based on Euclidean geometry (Bartholomew et al., 2008), as such it is often used as an explorative solution to learn about the associations between multiple variables in an analysis.

In our context, we opted for a cluster analysis over model-based approaches, such as latent class analysis, because we wanted to learn about the clustering process (how variables joined together in forming clusters, and how similar or dissimilar – strong or weak – these clusters could be considered). A model-based approach would have offered stronger statistical evidence on the robustness of the clusters (and on how many clusters can be identified in the data) but would have offered little information on how these clusters formed⁸.

In our analysis we proceeded with clustering variables rather than cases (the latter is the most common approach). We selected all the risk factors identified in the data and applied the cluster analysis to them. The correlation matrix, which forms the basis of the cluster analysis, was built using the Jaccard coefficient (J). This coefficient compares the presence of two binary items, looking at the ratio between their simultaneous presence, over all the possible combinations, but ignoring co-absences. For each pair of items, they can either co-exist (1,1), they can be both absent (0,0), or we can observe only the first (1,0) or only the second (0,1). The coefficient is based on how many times the following combinations can be found in the data:

$$J = \frac{(1,1)}{(1,1) + (0,1) + (1,0)}$$

This resulted in the Jaccard matrix in Table 11.1. For comparison the correlation between each item is also presented (Table 11.2). The Jaccard matrix was then analysed using a “complete” clustering process, based on the nearest neighbour methodology (at each clustering stage, the two clusters with the smallest dissimilarity are clustered together). The process was repeated with other clustering functions, but they returned consistent results with the “complete” solution.

The clusters were then extracted based on substantial considerations: at which stage of the clustering process we were able to identify clusters that were substantially relevant and seemed to describe a

⁸ One limitation of cluster analysis over model-based approaches is the impossibility to account for sampling error and survey nonresponse using survey weights in the clustering process. This means that the results of the analysis are not corrected for the under or over representation of some subgroups of the population.

meaningful set of clusters within the data. On this basis we set this threshold at a height of 0.85. This is on a 0-1 scale, where the height is a measure of similarity or dissimilarity between clusters (the smaller the height, the more similar are the clusters). Consequently, we can infer that the six clusters we identified in the data are quite weak, as they only clustered at the more dissimilar end of the scale – close to 1 – indicating that the clustered variables were still quite dissimilar.

Table 11.1 Jaccard matrix

Risk factors	Special educational needs	Lived in care/not with birth parent	Single parent household	Parental disability	Contact by education welfare services	Contact by police	Anti-social behaviours	Social renting	Absence from school	Expelled or suspended	Low academic attainment	Parental unemployment	Parent no qualifications	Low parental involvement in education	Depression	Child (pre-21)	Child (between ages 21-25)	Limiting disability	Caring responsibility
Special educational needs	-																		
Lived in care/not with birth parent	0.05	-																	
Single parent household	0.15	0.03	-																
Parental disability	0.13	0.03	0.12	-															
Contact by education welfare services	0.14	0.06	0.16	0.12	-														
Contact by police	0.11	0.04	0.12	0.09	0.22	-													
Anti-social behaviours	0.14	0.03	0.20	0.14	0.17	0.19	-												
Social renting	0.16	0.04	0.32	0.20	0.16	0.12	0.18	-											
Absence from school	0.10	0.05	0.11	0.09	0.20	0.13	0.11	0.13	-										
Expelled or suspended from school	0.15	0.05	0.14	0.10	0.26	0.24	0.20	0.16	0.14	-									
Low academic attainment	0.24	0.05	0.18	0.13	0.18	0.15	0.16	0.22	0.16	0.20	-								
Parental unemployment	0.07	0.02	0.09	0.11	0.06	0.05	0.09	0.17	0.07	0.07	0.09	-							
Parent has no qualifications	0.12	0.03	0.13	0.22	0.10	0.08	0.14	0.27	0.09	0.11	0.20	0.15	-						
Low parental involvement in education	0.08	0.02	0.10	0.07	0.07	0.08	0.09	0.09	0.08	0.08	0.08	0.05	0.07	-					
Whether depressed (most recent wave)	0.12	0.02	0.17	0.14	0.12	0.09	0.17	0.17	0.09	0.10	0.12	0.08	0.15	0.07	-				
Has own child (before age 21)	0.07	0.04	0.10	0.08	0.11	0.10	0.09	0.12	0.14	0.10	0.14	0.05	0.07	0.06	0.07	-			
Has own child (between ages 21-25)	0.11	0.03	0.15	0.11	0.11	0.10	0.13	0.18	0.11	0.11	0.18	0.08	0.14	0.07	0.11	0.24	-		
Limiting disability	0.21	0.03	0.14	0.14	0.16	0.10	0.13	0.14	0.13	0.11	0.16	0.07	0.10	0.06	0.21	0.07	0.09	-	
Has caring responsibilities post 16	0.16	0.03	0.22	0.19	0.13	0.10	0.21	0.25	0.10	0.13	0.17	0.11	0.27	0.08	0.20	0.13	0.23	0.14	-

Table 11.2 Correlations between risk factors

Risk factors	Special educational needs	Lived in care/not with birth parent	Single parent household	Parental disability	Contact by education welfare services	Contact by police	Anti-social behaviours	Social renting	Absence from school	Expelled or suspended	Low academic attainment	Parental unemployment	Parent no qualifications	Low parental involvement in education	Depression	Child (pre-21)	Child (between 21-25)	Limiting disability	Caring responsibility
Special educational needs	-																		
Lived in care/not with birth parent	0.07	-																	
Single parent household	0.06	0.02	-																
Parental disability	0.04	0.00	0.01	-															
Contact by education welfare services	0.12	0.11	0.14	0.06	-														
Contact by police	0.08	0.05	0.09	0.04	0.28	-													
Anti-social behaviours	0.04	0.03	0.10	0.03	0.17	0.24	-												
Social renting	0.07	0.07	0.30	0.15	0.14	0.10	0.07	-											
Absence from school	0.08	0.06	0.12	0.07	0.27	0.16	0.10	0.14	-										
Expelled or suspended from school	0.14	0.07	0.11	0.05	0.33	0.31	0.23	0.14	0.16	-									
Low academic attainment	0.21	0.04	0.10	0.05	0.15	0.15	0.09	0.16	0.11	0.18	-								
Parental unemployment	0.00	-0.00	0.01	0.06	0.01	0.00	0.02	0.19	0.05	0.03	0.05	-							
Parent has no qualifications	-0.01	0.03	-0.07	0.17	0.02	0.00	-0.03	0.21	0.06	0.04	0.12	0.16	-						
Low parental involvement in education	0.03	-0.01	0.08	0.01	0.03	0.06	0.05	0.05	0.07	0.05	0.06	0.00	0.01	-					
Whether depressed (most recent wave)	0.01	0.00	0.03	0.03	0.06	0.02	0.04	0.05	0.06	0.01	0.00	0.01	-0.00	0.01	-				
Has own child (before age 21)	0.04	0.03	0.11	0.05	0.12	0.11	0.07	0.15	0.18	0.11	0.13	0.02	0.04	0.03	0.02	-			
Has own child (between ages 21-25)	0.02	0.01	0.07	0.03	0.06	0.06	0.05	0.13	0.10	0.07	0.12	0.02	0.05	0.02	0.00	0.35	-		
Limiting disability	0.20	0.03	0.05	0.09	0.16	0.07	0.05	0.06	0.14	0.07	0.09	0.01	-0.02	0.01	0.19	0.05	0.01	-	
Has caring responsibilities post 16	0.02	0.04	0.05	0.08	0.06	0.05	0.05	0.13	0.07	0.07	0.10	0.07	0.11	0.02	0.04	0.23	0.24	0.03	-

Regression analysis

Binary logistic regression was used in this analysis because it is suitable for binary categorical outcomes. It was used to estimate the association between the different risk factors we identified and our outcome of interest – having been NEET at either Sweeps 6 or 7 or 8 of Next Steps.

Table 11.3 shows the results of the separate regression analyses run for each domain followed by the ‘all domains’ model’, which is the main model discussed in this report. This model has the lowest Log-Likelihood of any of the models, indicating that it has the best fit to our data. It also identifies several risk factors which ceased to be statistically significant once the role of all the risk factors are controlled for in a single model.

The results shown in Table 11.3 are the odds ratios. These reflect the change in the odds of being NEET if a risk factor is present, whilst controlling for other variables. For example, the first result in the model for the education domain indicates that the odds of young people with special educational needs becoming NEET are 1.865 times higher than the odds for young people who do not have special educational needs, controlling for other variables. In other words, a young person with special educational needs has 86.5% higher odds of becoming NEET compared to a young person without special educational needs, when keeping constant all the other elements included to the model.

The results for the regression analyses of the six clusters found in the cluster analysis can be found in Table 12.8 to Table 12.13. These show two stages of analysis conducted for each cluster. The first is a model including all the variables identified as belong to that cluster, with NEET status as the outcome variable. These consistently found each predictor to be a statistically significant predictor of NEET status. The second was the addition of interaction effects into the model. The term for the interaction effect, denoted in the tables as “variable 1 x variable 2”, indicates the effect on a person’s likelihood of being NEET of being present on both variable 1 and 2. In most cases, these were not statistically significant and did not improve the models fit. The one instance where an interaction effect was statistically significant, the interaction between being SEN and not holding a Level 2 academic qualification, can also be seen in Table 11.3. This shows that a young person who was **both** had SEN and did not hold a qualification at Level 2 was predicted to have a fall in their odds of about half compared to people who did not have either of these characteristics, in addition to the effects of these variables individually.

Table 11.3 Regression models

Dependent variable: NEET status						
Models	Education	Risky behaviour	Health	Living Standards	Family	All domains
Predictors and their odds ratios (95%CI)						
Special educational needs (SEN)	1.865*** (1.655, 2.075)	–	–	–	–	1.634*** (1.414, 1.854)
School absences	2.072*** (1.835, 2.308)	–	–	–	–	1.184*** (0.909, 1.460)

School suspension or exclusion	1.520*** (1.289, 1.750)	–	–	–	–	1.170*** (0.928, 1.411)
Academic qualification below Level 2	4.575*** (4.348, 4.801)	–	–	–	–	2.875*** (2.624, 3.127)
Low parental involvement in education	1.356*** (1.123, 1.589)	–	–	–	–	1.271*** (0.997, 1.544)
SEN × Academic qualification below Level 2	0.523*** (0.144, 0.903)	–	–	–	–	0.583*** (0.197, 0.969)
Services contact about YP	–	2.355*** (2.151, 2.560)	–	–	–	1.176*** (0.940, 1.411)
Police contact about YP	–	1.433*** (1.209, 1.657)	–	–	–	1.038*** (0.782, 1.294)
Anti-social behaviours	–	1.491*** (1.308, 1.674)	–	–	–	1.187*** (0.999, 1.375)
Mental health issues	–	–	1.564*** (1.408, 1.720)	–	–	1.656*** (1.475, 1.837)
Limiting disability	–	–	3.019*** (2.845, 3.194)	–	–	2.356*** (2.160, 2.551)
Social renting	–	–	–	3.193*** (3.031, 3.355)	–	1.467*** (1.274, 1.660)
Parental unemployment	–	–	–	1.425*** (1.190, 1.660)	–	1.360*** (1.087, 1.632)
Been in care/without natural parents	–	–	–	–	1.988*** (1.587, 2.389)	1.473*** (1.084, 1.861)
Single parent household	–	–	–	–	2.095*** (1.939, 2.252)	1.426*** (1.245, 1.608)
Parent has no qualifications	–	–	–	–	1.819*** (1.653, 1.984)	1.371*** (1.181, 1.561)
Parental disability	–	–	–	–	1.393*** (1.197, 1.589)	1.042*** (0.839, 1.244)
Has child (before age 21)	–	–	–	–	2.912*** (2.662, 3.162)	2.219*** (1.938, 2.501)
Has child (between ages 21-25)	–	–	–	–	1.914*** (1.729, 2.099)	1.827*** (1.619, 2.035)
Caring responsibilities (post 16)	–	–	–	–	1.390*** (1.227, 1.553)	1.329*** (1.158, 1.500)

Model statistics

Observations	7,209	7,274	7,635	7,667	7,434	7,082
Residual deviance	7,223.32	7,828.52	7,983.01	7,959.77	7,368.18	6,390.00
Log-Likelihood	-3,611.66	-3,914.26	-3,991.50	-3,979.89	-3,684.09	-3,195.02
McFadden's	0.09	0.02	0.05	0.06	0.10	0.17

Pseudo-R²

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

The NEET Risk Index

An index is defined as a formative measure (De Vellis and Thorpe, 2022; Diamantopoulos and Winklhofer 2001), meaning that the items that compose the index form the concept that the index is supposed to represent. In our case, several NEET risk factors form a general NEET risk index.

This marks a significant difference from reflective measures, such as scales, where the items are expected to be the reflection of a latent concept, and this can be extrapolated from the correlation of all the items. In our case, the correlation matrix of all the NEET risk factors in our analysis is expected to reveal multiple latent concepts which may reflect a condition of socio-economic deprivation, anti-social behaviour, or general poor health, but not a measure of NEET risk. In other terms, items composing an index are not expected to correlate – they may correlate, but their correlation may not necessarily reflect the concept of interest (in our case, NEET risk). This can lead to an analysis fallacy known as interpretational confounding (Howell et al. 2007), where the researcher mis-interprets the meaning of the measure they produced.

Due to these different correlation-related assumptions, the process of building an index cannot be based on methods based on classical response theory (such as factor analysis, used for scales), and other solutions should be undertaken.

In this appendix, we focus on how we address the elements below:

- **Redundancy.** Are we double counting index items, effectively inflating the role of some features in forming the index?
- **Weight.** What is the individual contribution of the different items in forming the index?
- **Validity.** Is the index measuring what we expect it to measure?

A regression-based approach

We addressed the first two points using a regression approach, commonly used for the development of a risk index (Zedler et al. 2015). This approach builds upon the formula of an index, defined as:

$$V = w_1I_1 + w_2I_2 + \dots + w_nI_n$$

In this formula, the index (V) is the sum of a number of items (I) multiplied for their respective weight (w). This is similar to a standard regression formula:

$$y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \dots + \beta_nx_n + \epsilon$$

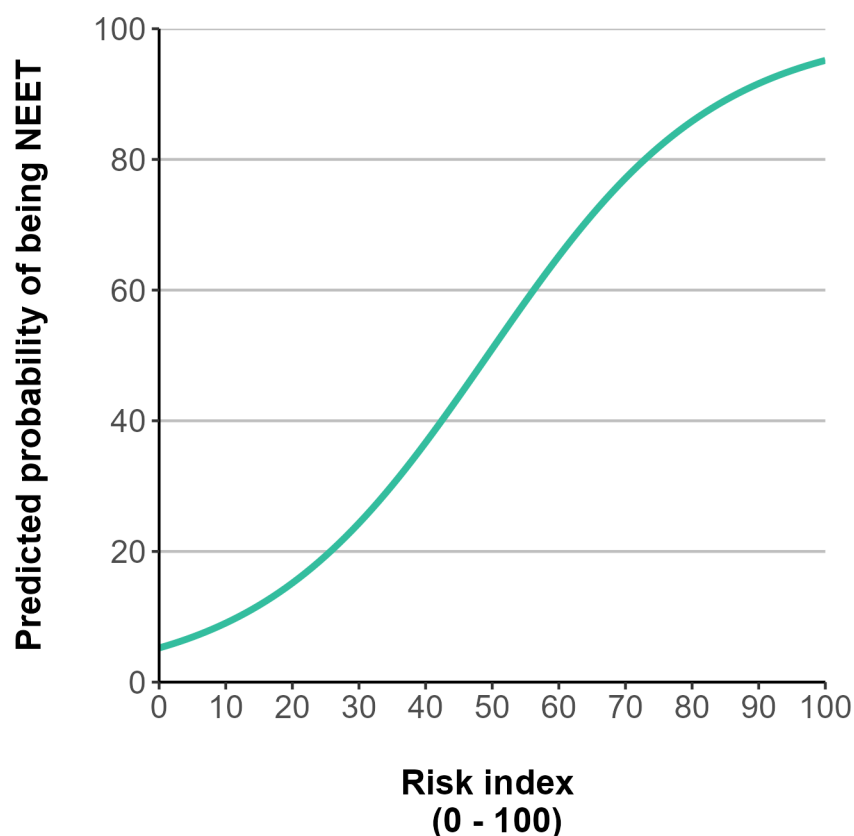
Where an outcome (y) is predicted by the sum of its predictors (x) multiplied by a coefficient (β), plus an intercept (β_0) and an error term (ϵ). This similarity between the two formulas has been used to identify the weight of the index item based on the regression coefficient ($w = \beta$). In other terms, by

fitting a regression to measure an observed outcome (y, in our context the presence of NEET status), it is possible to learn the weight of the different indicators and build a risk score for the observed outcome.

This addresses the first two elements listed above (redundancy and weight). Risk factors that have been “double counted” (association between predictors) are expected to result in smaller regression coefficients, offering an empirical solution to the redundancy problem. Similarly, the regression coefficients are expected to reflect the strength of the association between the predictor and the outcome: this means that index items that contribute the most to defining the outcome, are also the ones with the largest weight in the NEET risk index.

The log odds produced by the final regression model (with all indicators of marginalisation and the interaction effect between SEN and low academic attainment) were used to produce the index (log odds are the natural logarithm of the odds; they are listed in Table 12.7). Once the sum of the log odds for each case had been calculated, the results were then standardised to vary from between 0 and 100. This preserved the relative distance between individuals scores but put them on a more readily interpretable scale.

Figure 11.1 Predicted probability of being NEET by NEET Risk Index score



The regression approach also provided the basis for the validity of the index, given that it was built on a regression that was measuring the presence or absence of NEET status in the sample. Figure 11.1,

showing the association between the NEET Risk Index and the presence of NEET status, offers evidence of criterion validity. As someone's Risk Index score rises, their predicted probability of being NEET rises along with it, showing that the Index is an effective tool for measuring our concept of interest – a young person's risk of being NEET.

Construct validity (the theoretical relationship between the index and other non-NEET variables) can be inferred indirectly from the findings of Chapter 6 and Appendix 1. Indeed, the index is showing empirical associations that are consistent with what was found in other research, for example, analysis of the Labour Force Survey has previously shown similar patterns of NEET risk by ethnicity, with people of black ethnicity and Pakistani / Bangladeshi backgrounds at greater risk of being NEET, and those of Indian or 'other' Asian backgrounds least likely to be NEET (Powell, 2021).

Limitations

There are two strong limitations to this approach. The first limitation is related to the bias emerging from the selection of the items used to construct the index (other influential items may have been omitted, either because they are not measured in the Next Steps data or because they have not been identified as being risk factors for NEET status). This means that the risk score in the index is conditional to the content of the index itself. Although this is a common problem of formative measures, some authors suggested solutions based on structural equation modelling to identify the variance in the outcome of interest that is unexplained by the item selected for the index construction (Diamantopoulos et al., 2008; Bollen and Davis, 2009).

Another limitation concerns the reliability of the index. Although reliability is generally considered to be a feature of scales, and not necessarily of indexes, it is common practice to build the index on a sample and test its validity on different samples, showing that it has a similar empirical behaviour in different contexts (and not only in the contexts where it was constructed). The research objectives and scope of this project did not allow us to engage in these tests, but they are recommended for future uses of this NEET Risk Index.

12. Appendix 5: Supporting tables

The following tables are provided as supporting information on the analysis conducted.

Table 12.1 to Table 12.6 show summary statistics for the six clusters found by the cluster analysis. They show the different potential combinations of the risk factors making up each cluster, either having no characteristics from a cluster, having only one characteristic, having two characteristics, or (where clusters included three variables) having all three characteristics from the cluster. For each of these combinations they show the percentage of people that fell into that group, the percentage of people in that group who were NEET (at either Sweeps 6, 7 or 8) and the unweighted number of cases in that group.

Table 12.7 shows the results from the full regression model, including all risk factors, but it presents the coefficients as log odds instead of odds ratios. These are included for information because they were used to derive the NEET Risk Index, rather than the odds ratios.

Finally, Table 12.8 to Table 12.13 show the findings from the separate regression analyses run on the different clusters, described in the Regression Modelling section of Appendix 4.

Table 12.1 Summary table for Cluster 1

Cluster 1			
Expelled/suspended from school, contact from police services, contact from social or educational services			
Marginalisation Indicators	% with risk factors	% NEET	Unweighted base for % NEET
None	69	17.9	5624
Parents contacted by social services about YP's behaviour	6.6	35.7	424
Expelled or suspended from school	6.1	31.1	369
Parents contacted by police about YP's behaviour	5.1	27.3	304
Parents contacted by police about YP's behaviour & Parents contacted by social services about YP's behaviour & Expelled or suspended from school	4.5	49.6	179
Parents contacted by social services about YP's behaviour & Expelled or suspended from school	3.8	47	194
Parents contacted by police about YP's behaviour & Expelled or suspended from school	2.7	37.9	133
Parents contacted by police about YP's behaviour & Parents contacted by social services about YP's behaviour	2.2	35.5	113

Source: Next Steps Sweep 8 (age 25).

Table 12.2 Summary table for Cluster 2**Cluster 2**

Whether had child (before age 21) and whether had child (between ages 21-25)			
Marginalisation Indicators	% with risk factors	% NEET	Unweighted base for % NEET
None	76.8	18.2	6314
Has child (between ages 21-25)	13.2	32.8	862
Has child before age 21 & Has child (between ages 21-25)	6.2	63.1	336
Has Child (before age 21)	3.8	47.1	195

Source: Next Steps Sweep 8 (age 25).

Table 12.3 Summary table for Cluster 3**Cluster 3**

Single parent household, social renting and anti-social behaviour			
Marginalisation Indicators	% with risk factors	% NEET	Unweighted base for % NEET
None	40.4	13	3659
Anti-Social behaviours	13	19.2	1000
Single parent household	10.6	20.2	747
Single parent household & Social renting	10.3	38.7	612
Social renting	9.1	32.7	699
Anti-Social behaviours & Single parent household & Social renting	7.3	50.7	336
Anti-Social behaviours & Single parent household	5.3	31.3	319
Anti-Social behaviours & Social renting	4	43.2	244

Source: Next Steps Sweep 8 (age 25).

Table 12.4 Summary table for Cluster 4**Cluster 4**

Special educational needs and Academic qualifications below Level 2			
Marginalisation Indicators	% with risk factors	% NEET	Unweighted base for % NEET
None	61.7	14.4	5495
Special educational needs	13.8	24.5	907
Academic qualifications below Level 2	12.7	47.3	745
Academic qualifications below Level 2 & Special educational needs	11.8	48.7	519

Source: Next Steps Sweep 8 (age 25).

Table 12.5 Summary table for Cluster 5**Cluster 5**

Caring responsibilities (post-16), parent has no qualifications and parental disability			
Marginalisation Indicators	% with risk factors	% NEET	Unweighted base for % NEET
None	35.9	14.4	2778
Caring responsibilities (post 16)	24.2	26.4	1697
Parent has no qualifications	9.5	23.5	772
Parent has no qualifications & Caring responsibilities (post 16)	9.5	37.6	737
Parental disability	6.5	19	468
Parental disability & Caring responsibilities (post 16)	6.1	32	380
Parental disability & Parent has no qualifications & Caring responsibilities (post 16)	4.9	47.2	393
Parental disability & Parent has no qualifications	3.4	32.4	286

Source: Next Steps Sweep 8 (age 25).

Table 12.6 Summary table for Cluster 6**Cluster 6**

Limiting disability and Mental health issues			
Marginalisation Indicators	% with risk factors	% NEET	Unweighted base for % NEET
None	64.1	17.6	5099
Mental health issues	17.2	25.1	1353
Limiting disability	10.3	39.3	668
Mental health issues & Limiting disability	8.4	50.1	529

Source: Next Steps Sweep 8 (age 25).

Table 12.7 Full regression model (all domains) – log odds

Predictors	Log odds
Special educational needs (SEN)	0.49
School absences	0.17
School suspension or exclusion	0.16
Academic qualification below Level 2	1.06
Low parental involvement in education	0.24
SEN × Academic qualification below Level 2	-0.54
Services contact about YP	0.16
Police contact about YP	0.04
Anti-social behaviours	0.17
Mental health issues	0.50
Limiting disability	0.86
Social renting	0.38
Parental unemployment	0.31
Been in care/without natural parents	0.39
Single parent household	0.36
Parent has no qualifications	0.32
Parental disability	0.04
Has child (before age 21)	0.80
Has child (between ages 21-25)	0.60
Caring responsibilities (post 16)	0.28
Observations	7,082

Table 12.8 Regression summary for Cluster 1 variables

Models	Dependent variable: NEET status	
	Without interaction terms	With interaction terms
Predictors and their odds ratios (95%CI)		
Parents contacted by social services about YP's behaviour	2.087*** (1.881, 2.294)	2.543*** (2.283, 2.804)
Expelled or suspended from school	1.833*** (1.607, 2.060)	2.067*** (1.784, 2.350)
Parents contacted by police about YP's behaviour	1.343*** (1.114, 1.571)	1.717*** (1.400, 2.033)
Parents contacted by social services about YP's behaviour × Expelled or suspended from school	–	0.772 (0.271, 1.273)
Parents contacted by police about YP's behaviour × Parents contacted by social services about YP's behaviour	–	0.575 (0.003, 1.148)
Parents contacted by police about YP's behaviour × Expelled or suspended from school	–	0.786 (0.149, 1.422)
Parents contacted by police about YP's behaviour × Parents contacted by social services about YP's behaviour × Expelled or suspended from school	–	1.431 (0.478, 2.383)
Model statistics		
Observations	7,328	7,328
Residual deviance	7878.191	7866.761
Log-Likelihood	-3939.096	-3933.381
McFadden's Pseudo-R ²	0.02	0.02
LRT p-value		0.238

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 12.9 Regression summary for Cluster 2 variables

Dependent variable: NEET status		
Models	Without interaction terms	With interaction terms
Predictors and their odds ratios (95%CI)		
Has child before age 21	3.734*** (3.497, 3.970)	4.007*** (3.644, 4.369)
Has child (between ages 21-25)	2.135*** (1.954, 2.317)	2.192*** (1.996, 2.389)
Has child before age 21 × Has child (between ages 21-25)	–	0.874 (0.408, 1.340)
Model statistics		
Observations	7,688	7,688
Residual deviance	7918.884	7918.25
Log-Likelihood	-3959.442	-3959.125
McFadden's Pseudo-R ²	0.07	0.07
LRT p-value		0.366

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 12.10 Regression summary for Cluster 3 variables

Models	Dependent variable: NEET status	
	Without interaction terms	With interaction terms
Predictors and their odds ratios (95%CI)		
Single parent household	1.545*** (1.383, 1.707)	1.699*** (1.459, 1.938)
Social renting	2.827*** (2.657, 2.997)	3.254*** (3.007, 3.502)
Anti-Social behaviours	1.633*** (1.461, 1.805)	1.593*** (1.361, 1.826)
Single parent household × Social renting	–	0.767 (0.388, 1.147)
Anti-Social behaviours × Single parent household	–	1.130 (0.703, 1.558)
Anti-Social behaviours × Social renting	–	0.985 (0.558, 1.412)
Anti-Social behaviours × Single parent household × Social renting	–	0.917 (0.300, 1.535)
Model statistics		
Observations	7,601	7,601
Residual deviance	7766.38	7759.421
Log-Likelihood	-3883.19	-3879.71
McFadden's Pseudo-R ²	0.07	0.07
LRT p-value		0.388

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 12.11 Regression summary for Cluster 4 variables

Models	Dependent variable: NEET status	
	Without interaction terms	With interaction terms
Predictors and their odds ratios (95%CI)		
Academic qualifications below Level 2	4.263*** (4.092, 4.435)	5.337*** (5.125, 5.548)
Special educational needs	1.471*** (1.282, 1.659)	1.928*** (1.725, 2.131)
Academic qualifications below Level 2 × Special educational needs	–	0.548*** (0.185, 0.911)
Model statistics		
Observations	7,647	7,647
Residual deviance	7680.07	7656.919
Log-Likelihood	-3840.035	-3828.4595
McFadden's Pseudo-R ²	0.09	0.09
LRT p-value		0.001

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 12.12 Regression summary for Cluster 5 variables

Models	Dependent variable: NEET status	
	Without interaction terms	With interaction terms
Predictors and their odds ratios (95%CI)		
Caring responsibilities (post 16)	2.048*** (1.899, 2.197)	2.139*** (1.939, 2.338)
Parent has no qualifications	1.802*** (1.641, 1.962)	1.831*** (1.564, 2.098)
Parental disability	1.415*** (1.232, 1.599)	1.402*** (1.052, 1.753)
Caring responsibilities (post 16) × Parent has no qualifications	–	0.917 (0.549, 1.285)
Caring responsibilities (post 16) × Parental disability	–	0.935 (0.465, 1.405)
Parent has no qualifications × Parental disability	–	1.112 (0.595, 1.629)
Caring responsibilities (post 16) × Parent has no qualifications × Parental disability	–	1.019 (0.300, 1.738)
Model statistics		
Observations	7,494	7,494
Residual deviance	7975.256	7973.781
Log-Likelihood	-3987.628	-3986.8905
McFadden's Pseudo-R ²	0.03	0.3
LRT p-value		0.912

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 12.13 Regression summary for Cluster 6 variables

Dependent variable: NEET status		
Models	Without interaction terms	With interaction terms
Predictors and their odds ratios (95%CI)		
Limiting disability	3.019*** (2.845, 3.194)	3.039*** (2.795, 3.282)
Mental health issues	1.564*** (1.408, 1.720)	1.572*** (1.377, 1.767)
Mental health issues × Limiting disability	–	0.984 (0.615, 1.354)
Model statistics		
Observations	7,635	7,635
Residual deviance	7983.008	7982.993
Log-Likelihood	-3991.5	-3991.4965
McFadden's Pseudo-R ²	0.05	0.05
LRT p-value		0.489

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

13. Reference List

Addario, G., Bhagat, Y., Breedvelt, J., Crowley, J., Hudson, N., Khriakova, E., and Wilson, C. (2023). *Defining Marginalisation in Children and Young People: A Rapid Evidence Assessment*.

Allen, M. (2014) *Local action on health inequalities: Reducing the number of young people not in employment, education or training (NEET)*. Public Health England, London. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/356062/Review3_NEETs_health_inequalities.pdf.

Bartholomew, D. J., Steele, F., & Moustaki, I. (2008). *Analysis of multivariate social science data*. CRC press.

Bollen, K. A., & Davis, W. R. (2009). Causal indicator models: Identification, estimation, and testing. *Structural Equation Modeling: A Multidisciplinary Journal*, 16(3), 498-522.

DeVellis, R. F., & Thorpe, C. T. (2021). *Scale development: Theory and applications*. Sage publications.

Diamantopoulos, A. and Winklhofer, H.M. (2001) Index Construction with Formative Indicators: An Alternative to Scale Development. *Journal of Marketing Research*, 38, 269-277.
<http://dx.doi.org/10.1509/jmkr.38.2.269.18845>

Diamantopoulos, A., Riefler, P., & Roth, K. P. (2008). Advancing formative measurement models. *Journal of business research*, 61(12), 1203-1218.

DWP (2023) *Family Resources Survey: financial year 2021 to 2022*. Available at: <https://www.gov.uk/government/statistics/family-resources-survey-financial-year-2021-to-2022/family-resources-survey-financial-year-2021-to-2022>.

Equality and Human Rights Commission (2017) *Measurement framework for equality and human rights*. Available at: https://www.equalityhumanrights.com/sites/default/files/measurement-framework-interactive_pdf.pdf.

Feng, Z., Everington, D., Ralston, K., Dibben, C., Raab, G., & Graham, E. (2015). *Consequences, risk factors, and geography of young people not in education, employment or training (NEET)*. Scottish Government. Available at: <http://www.gov.scot/Publications/2015/10/2258/downloads#res487854>.

Holmes, C., Murphy, E. & Mayhew, K. (2021), "What accounts for changes in the chances of being NEET in the UK?", *Journal of education and work*, vol. 34, no. 4, pp. 389-413.

Howell, R. D., Breivik, E., & Wilcox, J. B. (2007). Reconsidering formative measurement. *Psychological Methods*, 12(2), 205–218. <https://doi.org/10.1037/1082-989X.12.2.205>

James, D., Yates, J. & Ferguson, E. 2013, "Can the 12-item General Health Questionnaire be used to identify medical students who might 'struggle' on the medical course? A prospective study on two cohorts", *BMC medical education*, vol. 13, no. 1, pp. 48-48.

Maychell, K., Evans, C., Brooks, R., Lee, B., and Pathak, S. (1998). Leaving at 16: a study of factors affecting young people's decision to leave full-time education. Available at: <https://www.nfer.ac.uk/media/1431/91107.pdf>.

ONS (2013) *Young people who are NEET: UK Estimate of Young People Not in Education, Employment or Training*. Office for National Statistics. Available at: https://webarchive.nationalarchives.gov.uk/ukgwa/20160105231310mp_/http://www.ons.gov.uk/ons/guide-method/method-quality/specific/labour-market/articles-and-reports/young-people-who-are-neet.pdf.

ONS (2016) *Young People Not in Education, Employment, or Training: February 2016*. Office for National Statistics. Available at: <https://www.ons.gov.uk/employmentandlabourmarket/peoplenotinwork/unemployment/bulletins/young-peoplenotineducationemploymentortrainingneet/february2016#definition-of-young-people-not-in-education-employment-or-training-neet>.

ONS (2023) *Young people not in education, employment or training (NEET), UK: May 2023*. Office for National Statistics. Available at: <https://www.ons.gov.uk/employmentandlabourmarket/peoplenotinwork/unemployment/bulletins/young-peoplenotineducationemploymentortrainingneet/may2023>.

ONS (2023b). Families in England and Wales: Census 2021. Available at: <https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/families/articles/familiesinenglandandwales/census2021>.

Public Health England (2021). Health-related quality of life for people aged 65 and over. Available at: <https://www.ethnicity-facts-figures.service.gov.uk/health/physical-health/health-related-quality-of-life-for-people-aged-65-and-over/latest>

Powell, A. (2021) *NEET: Young people Not in Education, Employment or Training*. House of Commons Library. Available at: <https://researchbriefings.files.parliament.uk/documents/SN06705/SN06705.pdf>.

Zedler, B., Xie, L., Wang, L., Joyce, A., Vick, C., Brigham, J., ... & Murrelle, L. (2015). Development of a risk index for serious prescription opioid-induced respiratory depression or overdose in Veterans' Health Administration patients. *Pain Medicine*, 16(8), 1566-1579.

