Predicting adult life outcomes from earlier signals: modelling pathways through childhood

Report for HMT, version 1.3

Leon Feinstein

Oct 16, 2006

Centre for Research on the Wider Benefits of Learning

Institute of Education, University of London

Executive Summary

Overview

This report provides an assessment of the extent of persistence and change in childhood risk of adult outcomes as children pass through childhood and adolescence. We assess the levels of change in risk status and model the implications of this for the assessment of the value and cost of different intervention scenarios.

The starting point for the analysis is the objective of assessing the practicality of early preventative intervention to reduce adult deprivation, childhood risk and the inter-generational persistence of inequality and disadvantage. The policy framework for the analysis is that of progressive universalism, i.e. the objective of providing support and intervention on the basis of need within a universal system recognising the entitlement of all to such support.

It is also important to emphasise that although it is possible to identify childhood risk and to predict likely outcomes for groups at risk, on average, it is never possible to predict the outcome for any individual child. One can make a forecast of likely outcomes and use this forecast to assess the cost-effectiveness and social value of a set of possible interventions but this is not the same as suggesting that the forecast for any individual child is set in stone. That would not only deny the possibility of intervention but also lead to stigmatising effects which would undermine the primary purpose of the intervention.

The ideal of progressive universalism is useful in this policy environment in part because it provides a basis for developing intervention designs which do not stigmatise those whose need justifies extra support. Any system of early intervention runs the risk of stigmatising those in receipt of the intervention but this must be minimised and seen in the light of the alternative risk, that of ignoring need, risk persistence and inequality and so allowing the propagation of disadvantage with all the resulting personal and social costs that brings.

In this report we assess the extent to which children move in and out of risk through childhood and model the implications of these movements for the cost effectiveness of different scenarios of intervention.

Key findings

The adult implications of childhood risk

The odds of experiencing the adult outcomes are very high for those who are defined as at risk at each age. Being at risk in childhood is very strongly associated with the likelihood of experiencing the indicated outcome.

Substantial odds of adult deprivation are observed to follow from risk status at birth. This is especially so for the 1970 Cohort. For example, for the multiple deprivation outcome, the true and false positive rates for the 1958 Cohort (from birth to age 23) are 37% and 8%; for the 1970 Cohort (from birth to age 30) the equivalent statistics are 52% and 9%. This is a big rise in the true positive rate without an equivalent rise in the false positive rate. These findings are mirrored across all outcomes.

Risk continuity
Being classified as at risk at age 10/11 adds substantially to the already high risk of experiencing adult deprivation that follows at-risk status at birth.

For many outcomes in the 1958 cohort, risk status at age 7 is a substantial augmentation to risk to that of birth. Age 5 risk in the 1970 Cohort, however, is for most outcomes, closer to birth risk than to age 10 risk in the strength of its negative implications for adult outcomes.

Risk continuity between sweeps at birth and age 16 is not much lower than that between birth and age 5/7. In some cases it is even higher. This reflects the impact of compounding risk. A fairly high proportion of those who exited risk status between birth and later ages, re-enter subsequently.

Overall, we find considerable mobility. Most of those at risk at one age are often not at risk at subsequent ages.

There is also important continuity in risk. The chances of being at risk at any age are much higher for those at risk at previous ages.

Overall, the findings here support the conclusion of a balanced perspective on risk continuity and discontinuity. Risk is not set in stone and immutable, neither is it random and unimportant for future chances.

Calibration of intervention scenarios
We distinguish between

- i) the direct benefit that results from a reduction in the likelihood of adult deprivation for those in receipt of intervention in childhood, and;
- ii) the indirect benefit that follows from the reduction in the need and cost of subsequent intervention.

Making the assumption that the cost and effectiveness of intervention is the same at all ages, we find that whereas the direct benefit is higher for intervention later in childhood, the indirect benefit is higher for earlier intervention. However, a strongly declining (high early intervention) schedule does not derive great cost savings from the reductions in later risk because in such a schedule relatively few at risk young people would in any case receive late intervention.

Conclusion
Integrating these findings into a single metric of the benefit to cost ratio, we conclude that the most effective intervention schedule is one which starts at birth and continues through childhood but at a very slightly decreasing rate. This maximises the benefits of reducing later need for intervention while also enabling the intervention design to pick up and support the many children whose need was not apparent at earlier ages.

This overall finding does not take into account the possibility that later intervention may be cheaper or easier to implement, as we assume constant cost. We also do not model the possibility that earlier intervention may be more effective. The possibility of greater returns for repeated intervention should also be allowed for as should the likelihood of diminishing returns to scale (universality).

The calibration framework applied here is, therefore, rather basic and we recommend that more sophisticated modelling be undertaken to determine the appropriate level and timing of intervention through childhood in a model which can address these important concerns.

However, we are confident that the general principle of the value of persistent intervention would emerge equally strongly in a more general and sophisticated calibration exercise

Table of Contents

1. Research focus and objectives

1.1 Building on previous research

This report builds on recent work conducted by the WBL for the PMSU on "*Predicting adult life outcomes from earlier signals: identifying those at risk,*" and extends that work with the aim of providing useful analytical information for the CYP Review in HMT.

The research conducted by the WBL for the PMSU focused on predicting adult outcomes from childhood information about family background and child development. It has been possible to show that those most at risk of outcomes of adult social exclusion can be identified fairly well, fairly early in childhood, although there are of course debates about the level of identification required to justify such a statement. The study indicated that preventative intervention targeted on those most at risk will tend to be highly cost effective, under reasonable assumptions about programme effectiveness and the cost of social exclusion. In other words, we found that the level of identification was sufficient to suggest that inability to identify those at risk was not a substantive barrier to the policy objective of preventative intervention.

We noted, however, that the relationship between childhood risk and high cost or high harm outcomes in adolescence or adulthood is not deterministic, mechanistic or inevitable. There are many steps on the pathway from risk to outcome. There are children at risk who do not experience harmful outcomes and there are children with low apparent or observable risk who do.

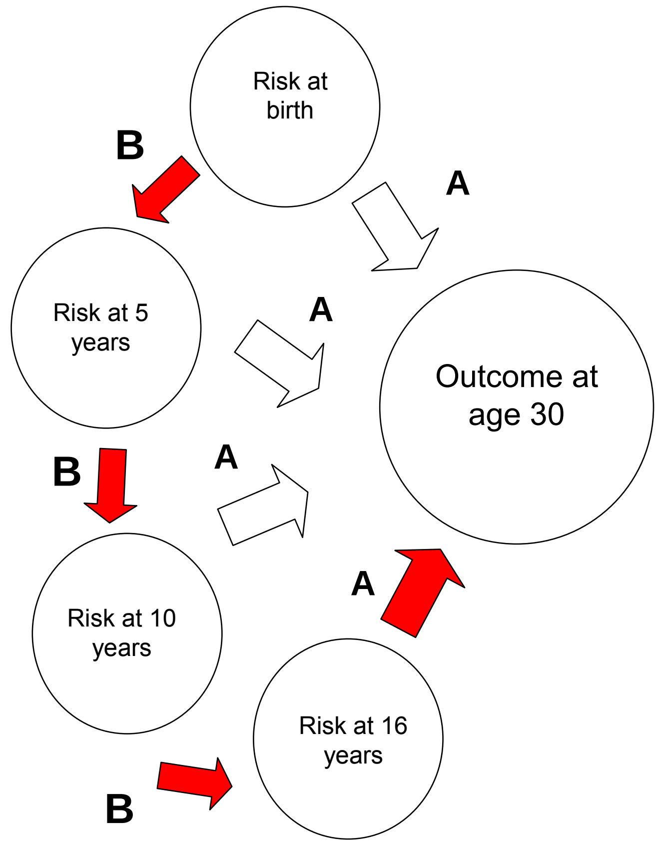
Therefore policy responses must allow for flexibility and change. Administrative data should always be augmented by local level, practitioner knowledge and the appropriate interventions should also be selected by local level practitioners who should work closely alongside local communities and agencies to avoid rigid tracking or excessive and unnecessary stigmatisation of vulnerable young people and their families.

In this current report, we explore in more detail some of these childhood pathways of movement in and out of risk and evaluate the implications of these continuities and discontinuities for the process of identification and intervention for those at risk, in a system based on the three key principles set out in the HMT/DfES review (HMT 2005), namely:

- rights and responsibilities: supporting parents to meet their responsibilities to their children;
- progressive universalism: support for all, with more support for those who need it most; and
- prevention: working to prevent poor outcomes for children, young people and their parents from developing in the first place.

Relevant risk pathways are described graphically in Figure 1, below. The static picture of the predictive capacity of risk at any one age for a particular adult outcome at age 30 is shown as pathway A in Figure 1. Some of the more dynamic potential risk pathways are indicated by pathway B.

Figure 1: Static and dynamic models of risk



Analysing pathway A only allows us to make judgements about the cost effectiveness and targeting of individuals interventions. For example, is Tots and Teens a cost effective programme to reduce the incidence of teenage pregnancy if targeted at the 5 percent most at risk at age 10? It does not tell us about the overall impact on adult outcomes of different interventions at different ages. This requires a more dynamic model as illustrated by pathway B. Analysis of pathways such as pathway B can inform the overall package of policies and interventions that were provided for children, young people and families. But there are many other potential pathways in and out of risk through childhood. These are considered in this report.

1.2 Overview of research questions

More specifically, we address two particular research questions:

1. Risk continuity. Children move in and out of risk. We define risk in terms of a high likelihood of experiencing adult deprivation and assess this risk at each age in childhood at which data is collected in the surveys used. Using a number of different approaches, we then assess the degree of continuity and discontinuity of risk and describe the pathways observed through childhood.

2. The calibration of intervention effects. Because children move in and out of risk, intervention designs which are static or fixed at a single age will miss many cases of need and fail to maximise the potential of preventative intervention. We evaluate different intervention scenarios which allocate preventative support at different rates at different ages to those identified as at risk.

2. Methods and research resources

2.1 Data

We draw on two datasets: the 1958 and 1970 UK birth cohorts:

National Child Development Study (a.k.a. 1958 Cohort)

The sampling frame for the NCDS was the population of all children born in Great Britain in the first week of March. Data was collected at birth and then again at ages 7, 11, 16, 23, 33 and 42. During childhood interviews were conducted with parents, teachers and medical officers and since 16, the sample members themselves have responded to extensive interviews and questionnaires.

British Cohort Study (a.k.a. 1970 Cohort)

The BCS has a similar structure to that of the NCDS with sampling in the first week of April 1970 and data collected at ages 0, 5, 10, 16, 26 and 30. Again, there has been extensive data collection by interview and questionnaire with the sample members, parents, teachers, health visitors and medical officers. Testing of cognitive development has also been undertaken.

The ’58 and ’70 cohorts have the strength that they have adult outcomes and also multiple measures through childhood from various sources. We describe the measures used in more detail below.

2.2 Signalling of risk is not causality

Under a model of progressive universalism in policy delivery, it is important to identify those with greatest need at the earliest possible opportunity and to provide appropriate support. If one recognises the objective of minimising the social costs that result from social exclusion and related outcomes of low productivity, social division and inter-generational inequity, then this requires identification of the relative risk of these outcomes for individuals, households and communities so as to target support. A great deal of intervention has been targeted at the community level but because of important within-community differences in need and access to resources, often the individuals and families most in need are missed by such an identification strategy. The ideal model would be one that recognised each of these units of analysis and used all in the identification strategy. The contribution of this research project, based as it is on individual level birth cohort data, is particularly focused on individual level targeting and the identification of those most “at risk.”.

It is important to emphasise that this information is not by itself sufficient for an accurate assessment of the value of intervention or of the relative effects of interventions at different ages. That assessment would require knowledge in relation to three types of question:

1. what are the factors that indicate risk (signalling question)?
2. what are the related causal mechanisms (causal question)?
3. to what extent is each mechanism amenable to policy intervention (policy effectiveness question)?

The signalling question

This report focuses on the first of these questions. The key research objective has been to assess and describe the extent to which accurate predictions can be made during childhood about who will experience various forms of deprivation later in adolescence and adulthood.

The precise level of accurate identification required to make intervention cost-effective depends on the relative costs of intervention and failure. There is no a priori minimum proportion of accurate forecasts that is required to suggest that identification and intervention is cost-effective. One might assess the rate of accurate identification of children at risk ("hit rate" or "sensitivity") for a particular outcome to be as low as, say, 30% but if the cost of failure to intervene and the success probability of intervention are both high then intervention may nonetheless be cost-effective. The required minimum hit rate depends on the trade-off of costs and can only meaningfully be identified on the basis of an assessment of this trade-off.

These cost/benefit trade-offs are modelled in previous work (Feinstein and Sabates[1]) under a range of assumptions about the social cost of outcomes, programme effectiveness and the costs of intervention. The empirical analysis of the identification of at-risk individuals is fed into a model to assess the cost-benefit ratios of specific interventions in different scenarios. This information is useful in providing a framework for the assessment of the social value of any specific intervention targeted at the prevention of adult deprivation.

That previous report also provides a description and application of a method (based on Receiver Operating Characteristics (ROC) curves) for assessing the appropriate level of universality or targeting of a specific intervention.

The causality question

We have undertaken this work within a coherent framework of prediction that clarifies the relationship of the measures used to the outcomes described. In other words this is not a simple data-mining exercise because it is important to found the discussion of the predictive power of the different measures on clear theoretical foundations that consider issues of causality and process. However, we do not in any way attempt to resolve these issues and the inferences drawn about identification and persistence of risk should not be equated with statements about causal mechanisms. Strong simplifying assumptions are used to assess the effectiveness of interventions under the different scenarios considered.

The policy question

This project has not assessed the feasibility and effectiveness of interventions, instead providing information that can be used in a model to ascertain the cost-effectiveness of intervention under a range of assumptions about intervention effectiveness.

---

[1] Feinstein, L. and Sabates, R. (Aug, 2006) ."Predicting adult life outcomes from earlier signals: identifying those at risk." Report for the Prime Minister's Strategy Unit. CLS Occasional Paper, Institute of Education, 2006.

Nonetheless, the specification of this signalling or identification question does raise issues for intervention design. It is important that the policy response fit with and be responsive to the identification process. The identification of an individual at risk should not necessarily trigger instant costly intervention. A well-designed and intelligent system of risk management would recognise the need to monitor levels of risk over time and aim to intervene at the time of greatest effectiveness. The identification of this time and of the appropriate intervention may not be best judged only from the perspective of administrative or quantitative data of the kind being used here but is most effective if also drawing on the skill, judgement and experience of a lead professional who knows the child and child's contexts.

This extra layer between identification of risk and subsequent action is vital if stigmatising effects and badly targeted interventions are to be minimised

Risks of identification

There are costs associated with excessively rigid targeting schemes that risk stigmatising children and damaging their self-concepts, thus adding to their level of risk. It is also helpful to recognise that the assessment of true or accurate forecasts for individuals will change over childhood as children's balance of risk and protection interacts with their own development to change the nature of the forecast for them. Thus an individual identified as high risk at one age will not necessarily be high risk at another age and there may be a substantial discontinuity of this kind and vice versa in terms of those becoming at risk.

2.3 Defining "at-risk" groups.

At each age at which data are collected in the study datasets, we can ascertain the likelihood for each individual in the data of experiencing each feature of adult deprivation. For each individual we have self-reported information on whether or not they did in fact experience the outcome and we can run statistical regression analysis to test the extent to which each observed feature of childhood development and context is on average associated with the adult outcome. This regression equation can also then be used as a forecast or prediction model, to create for each individual a likelihood of experiencing the outcome given their balance of risks and opportunities in the observed childhood data. Children with the highest level of likelihood of experiencing the adult outcome can in these terms be defined as "at risk."

We can then compare the actual outcomes to those forecast, i.e. the assessment of the rates of true and false positives in the data, terms defined shortly, below. This risk forecast can be developed using data at particular ages, from particular sources and in terms of particular domains of childhood risk. In this way, the implications for adult outcomes of signals from these different ages, sources and domains can be tested and compared.

However, the precise cut-off at which an individual is said to be "at risk," as opposed to not "at risk," is an arbitrary and to some extent false dichotomy, as rather what is observed is more likely to be a continuum of risk. More sophisticated approaches (e.g. Swets, Dawes, and Monahan[2]; Billings, Dixon, Mijanovich, Wennberg [3]) model this continuum and test the implications for predictive capability of different cut-off points, assessing the accuracy of predictions through summarising accuracy across the full range of potential cut-off points.

Here, because of the wide range of outcomes and time periods considered, we work with a single arbitrary cut-off, namely that which leads to a prediction of risk status which has the same prevalence as observed for the outcome. In other words, if an adult outcome has a prevalence of 20%, we create a risk group that has the same rate of prevalence. If such arbitrary risk groups were used as the sole method of policy targeting, this would lead to inefficient and ineffective allocation of interventions. It is not at all likely that an effective intervention would be of benefit to those in the most at risk 20% of the population but not to individuals at only marginally lower risk. The full balance of trade-offs of costs and benefits requires a more complex and sophisticated modelling exercise (see Feinstein and Sabates, *op. cit.*, for a more formal discussion) and it may be appropriate to target interventions at a higher or lower proportion than that observed to in fact experience the outcome of concern.

The risk groups created here are not constructed in order to target intervention as part of a real system of policy delivery but to assess the relationships between being at risk at one age and at risk at another. The method has the virtue of relative simplicity and easy presentation. Under strong linearity assumptions, the cut-off issue need not be a concern but should be remembered as a caveat in the interpretation and understanding of the findings.

2.4 True and False Positives

The simplest model takes the form of a linear prediction equation, with each relevant adult outcome treated as a separate dependent variable. The sets of risk and protective factors are entered on the right hand side and the findings indicate the extent to which risk and protective factors at different ages enhance predictive capability relative to false predictions.

There are a number of ways of assessing the predictive power of measures of risk and protection. It is common in statistical analysis to assess the quality of the prediction in terms of the proportion of variance in an outcome explained by the predicting measures, or in terms of the statistical significance of predictor variables. To focus on the issue of the identification of individuals at risk, we use an alternative approach.

---

[2] Swets, J. A., Dawes, R. M. and Monahan, J. (2000). "Psychological Science Can Improve Diagnostic Decisions." Psychological Sciences in the Public Interest (a supplement to Psychological Science), 1(1)"

[3] John Billings, Jennifer Dixon, Tod Mijanovich, David Wennberg, (2006). "Case finding for patients at risk of readmission to hospital: development of algorithm to identify high risk patients." BMJ 2006;333:327

The focus of the analysis here is on the quality of the identification of those at risk of adult deprivation and so being potentially triggering a prior preventative intervention or interventions. The measures of childhood circumstances and child development are used to predict later outcomes. This generates a statistical model that effectively weights each childhood construct in terms of its importance as a signal of the risk that the outcome will occur. These weights are then linked to the observed childhood data for sample members to generate for each sample member a probability that they will experience the adult outcome. Much of the analysis then focuses on this probability which can also be termed a risk propensity score.

For each individual in the data, the data and analysis provides an observation of whether or not they actually experienced the outcome of adult deprivation (truth) and a prediction based on their earlier childhood risk factors of their likelihood of experiencing the outcome (prediction).

This approach creates a two-way table of probabilities, as shown in Table 1.

Table 1: A tabular description of forecast accuracy for binary outcome.

		Truth		
		Positive (T+)	**Negative (T-)**	
Decision (Predicted outcome)	**Positive (D+)**	True positive A	False positive C	P(D+) = A+C
	Negative (D-)	False negative B	True negative D	P(D-) = B+D
Note: A+B+C+D=1		P(T+) = A+B	P(T-) = C+D	

Two proportions provide a very good indication of the predictive capability of the measures that were used to generate the predicted outcome, namely:

1. *The true positive proportion P(TP)*, the proportion of those who in actuality experienced the outcome who are accurately predicted to do so by the risk propensity score[4];
2. *The false positive proportion*, P(FP) i.e. the proportion of those who did not experience the outcome who are erroneously predicted to do so by the risk propensity score;

Note that as defined here:

[4] It is important to accurately specify the definition of true and false positives in any study as these terms are sometimes defined as here, i.e. conditional on the true state, but can also be defined in terms of the probability of the true state conditional on the prediction. This alternative formulation of P (TP) is as the proportion of predictions that are accurate. Which formulation is to be preferred depends on the use to which the information is to be put.

$P(TP) = A / (A+B) = P(D+,T+) / P(T+) = P(D+ | T+)$

$P(FP) = C / (C+D) = P(D+, T-) / P(T-) = P(D+ | T-)$

These two proportions provide an indication of the accuracy of the identification of those at risk of adult deprivation.

As Table 1 highlights, the usefulness of the prediction information provided by this study can be in assessing the relative likelihood for a range of outcomes of:
i) True negatives - accurate identification of those not requiring extra support;
ii) False positives - deadweight intervention;
iii) False negatives - missed intervention needs resulting in failure to intervene;
iv) True positives - accurately targeted interventions.

There are costs, savings and/or benefits associated with each type of assessment status.

It is important to note that as the threshold determining the prediction changes, there will be important implications for the proportions of true and false positives. A low threshold which creates a prediction that most individuals will be predicted to experience the outcome is likely to accurately identify a large proportion of those who eventually do have the outcome (a high true positive rate) but also to falsely suggest that the outcome will be experienced by a lot of those who did not in reality experience the outcome (a high false positive rate). Thus, the decision rule which sets this threshold and determines the proportion of those who are to be classified as at risk or positive predictions is a crucial factor in more sophisticated models.

2.5 Treatment of missing values and missing cases

In this report we assess the accuracy and continuity of prediction when a great many measures during childhood are used to predict adult outcomes. An important methodological issue is how cases are treated when information on a particular measure is missing for particular individuals, but when information is available in relation to other measures. It is standard practice to maintain a high sample size in such instances by using mean imputation, multiple imputation or other forms of data substitution. Assuming that missingness is at random, conditional on other data in the model, these methods do not cause bias, although standard errors must be corrected.

For these predictive models, the impact of missingness is rather different. If we drop cases with missing data on any single measure then the number of observations maintained in these very rich regression models becomes very low. However, these are cases for which we have the best data and, therefore, for whom the accuracy of prediction will be highest because we have the most information. These results indicate the potential of reliable and high response data to identify children at risk. However, equivalent levels of identification would require that the data were collected without high levels of missingness and attrition, as might in certain circumstances be easier with administrative data collection than with the longitudinal surveys used here. Assuming that cases with full data are not more predictable than other cases, there is no bias induced to the estimate of the true and false positive proportions by generalising from these small samples. However, a degree of sampling bias may result in any small sample analysis and inference becomes more hazardous.

The average level of adult deprivation will be lower in the full information data than in the full sample because it tends to be those most at risk on various indicators who are most likely to leave surveys or to fail to respond to specific questions. Thus the full information samples include fewer cases of multiple deprivation than is the case for the full sample.

Therefore, as well as reporting results when only the full information cases are used, we also report, for the BCS and NCDS, results when mean imputation methods are used. These results are based on far larger samples but include cases for whom less information is available. This makes prediction less accurate.

Taken together, the two different sets of results perhaps provide upper and lower bounds as to the degree of early identification possible. Providing both sets of estimates enables an assessment of this range of possibilities. It would be possible to experiment with more efficient forms of data imputation to test the levels of prediction in larger maintained samples at the later ages.

3. The prediction data and outcomes

There are a great many features of the background and development of children that may provide signals about likely future outcomes. In order to structure the analysis and to provide a foundation for interpreting the results we provide a suggested theoretical framework.

3.1 Theoretical framework

The general theoretical framework for considering these relationships is derived from the ecological model of Bronfenbrenner[5], together with more recent developing science in bio-medical, psychological social science fields. We emphasise the following key principals of human development, particularly during the period of childhood and adolescence:

i) There are features of biology and temperament that are determined by inherited, genotypic characteristics and which have important implications for subsequent development and outcomes;

ii) These genetic features interact with features of the environment that therefore also have important implications for development and outcomes;

iii) There are many features of the environment that may impact on human development to influence the relative likelihood of adult outcomes for different children but key amongst these influences, particularly but not exclusively in the early years are those experienced in the family;

iv) Other contexts also matter, increasingly through childhood. Particularly important other contexts are peer groups, schools, neighbourhood factors and wider social and community networks. As children move through childhood into adulthood, a wider array of contexts start to play an increasingly important role, particularly factors such as labour markets, work and college environments, close personal relationships and other interactions with adults in positions of authority and influence;

v) These contexts interact in their influence in the development of children and young people;

vi) The personal skills, attributes, characteristics, temperaments and self-concepts of children and young people also contribute to their success in negotiating the pathway through childhood into adulthood. Children have some degree of agency in their own development but many of the underlying attributes which support agency will also be strongly influenced by background and contextual features.

vii) Thus development occurs through a multi-faceted and complex series of interactions between the developing child and the other agents and agencies with which the child forms relationships;

viii) Pathways are not set in stone as a result of early development or contexts and there is always the possibility of discontinuity or plasticity which can change likely outcomes. However, disadvantage and background stress exerts continuous and

[5] E.g., Bronfenbrenner, U. (1979). *The ecology of human development.* Cambridge, MA: Harvard University Press.

compounding pressures on young people such that the discontinuites in development tend to privilege those from more advantaged backgrounds, i.e. with more supportive, nurturing and well-resourced environments.

These principles are intended to clarify the analysis undertaken on the signals of risk. It should be clear that not all of the factors that influence development can be measured in standard data collection exercises. Many of the important causal mechanisms are the result of very rapid intra-personal neurological and biological processes that cannot commonly be observed and which interact with features of the developing child and of social and family contexts in dynamic transactions which are not yet well understood. Therefore, it would be wrong to expect that these causal processes can be easily manipulated in a theoretically grounded science of policy delivery.

However, as a result of these causal processes there are relatively stable features of personal development and of social and family context which indicate the general trends in development and carry strong signals about likely outcomes. These signals are constructs which can be measured and modelled.

3.2 Domains of measurement

The features of contexts that are supportive of positive development can be classified in terms of their relationship to the developing child, namely:
Proximal processes are the features of the interaction between the child and others in the family, typically the primary care giver. The proximal processes, such as home learning, are child-centred interactions, reciprocal, dynamic and transactional;
Distal family factors refer here to observable features of the members of the family or of the general family context as a unit that impact and influence proximal processes, as well as having direct impacts on child development.

Seven domains of measurement

This framework provides us with 7 basic classifications of variables or domains of measurement, namely:

Family context
1) Distal family factors
2) Proximal family process

Other contexts
3) Neighbourhood factors
4) School factors

Child factors
5) Features of the cognitive development of the child
6) Features of the affective and behavioural development of the child
7) Child health

Most of the measures used to predict outcomes in adulthood and adolescence can be classified into one of these 7 domains. Other datasets also include data on a wider range of contexts in which case the list of domains can be extended. Moreover, some datasets might provide the capability to distinguish between distal and proximal features of contexts other than the family, but that is not the case with the datasets used in this study and so the proximal/distal distinction is only maintained here for the family context.

There are some complex issues in the classification of constructs that are not always straight-forward to resolve. For example, if a child attends an independent school is that a feature of the school context or more accurately of the family context as it is indicative of family characteristics such as wealth and aspiration as well as of the likely characteristics of the school. Similarly, if a young person spends a lot of time in the company of peers engaged in activities that might be classified as anti-social, does this provide a signal about the peer group, properly categorised as part of the peer context, a separate domain of measurement, or is that more indicative of the affective and behavioural development of the young person?

These issues cannot be simply resolved and result from the obliqueness, multiplicity and inter-connectedness of the underlying causal mechanisms. Our strategy in this report is to be transparent about the ways in which measures are classified so that inferences can be appropriately discussed as required.

Commensurability of measurement

The datasets used for this study are high quality, longitudinal datasets that are internationally renowned as being a major scientific resource of a kind not available in such depth and breadth anywhere else in the world. They are the result of long-term investments by scientific research councils, academics and sponsoring Government departments and make possible an in-depth and detailed analysis of the research questions of concern in this report.

However, it is nonetheless important to recognise the limitations of these datasets and the implications of those limitations for inference in relation to the findings of this research.

One particularly important issue is that for a number of reasons comparability across ages within studies and between studies must be undertaken with considerable care. The measures were collected as part of separate studies that were in differing amounts and at different times, differently intended to meet the different and changing research objectives of bio-medical science, social science and policy. The 1958 Cohort Study (NCDS) was originally intended to support analysis of perinatal mortality and was initially developed on the basis of a research design appropriate for that study. The 1970 Cohort Study (1970 Cohort) suffered from the co-existence of a teachers' strike with the age 16 data collection that led to big problems of missing data. The datasets have used different forms of data collection at different ages of the study children. The main focus of the 1958 and 1970 cohorts was initially on health and development and there are few available measures of the details of family life.

The precise forms of data collection in each study by age are set out in Table 2.

Table 2: Sources and ages of data collection in the NCDS, BCS

		Age (years)																
	Source	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1958 cohort	Teacher								√				√					√
	Parent	√							√				√					√
	Medical												√					√
	Child												√					√
1970 cohort	Teacher											√						√
	Parent	√					√					√						
	Medical											√						
	Child											√						√

Because of the differences in focus of the study at different ages the quality of measurement across domains and ages varies quite considerably such that there is not commensurability of measurement across the 7 domains. Measurement of the neighbourhood context is particularly weak, there being few easily available measures collected in the data and the relevant constructs being less well understood and possibly harder to measure in individual level data than those for the family context. Thus scientific comparison of the relative importance of these different constructs as indicators of development must be undertaken with caution.

The purpose of the analysis presented here is to assess the differences in relative predictive power of the measures at different ages rather than across domains. Since there is differential assessment at each age in terms of the range of measures assessed, some of the change in risk status will result from change in the format and range of measurement, not necessarily from a real shift in the risk for children. This should be borne in mind.

Accessibility of measurement

It is also important to emphasise that the capability of survey agencies to collect data as part of a scientific study may be different in important ways to that of Government agencies engaged in policy. Government-led data collection may, subject to legal and administrative issues, have good access to administrative data routinely collected by education, social services and health departments and agencies, although there remain barriers to the linkage of such data and issues of anonymity, ethics and access to the data.

The legislative power of Government and its ability to regulate and create frameworks for assessment and monitoring may change the context within which important constructs are to be measured such that the data collected as part of the policy framework may be more accurate and suffer less from problems of attrition and missing values than do the data analysed in this study. To the extent that is the case, the results from this study provide a lower bound or under-estimate of the true predictive power and possibly an overestimate of persistence that might be generated in a policy-led system of data collection.

However, it may also be the case that for some sources of data collection such as parents and children the access to information may be easier for a disinterested survey design company than it is for government agencies, particularly if there may be punitive or otherwise unwanted repercussions resulting from particular assessments. This may result in mis-measurement or missing values, perhaps particularly for the most important cases if the data collection exercise is not handled sensitively and if the policy system that the data collection informs is seen as a threat rather than an opportunity. Balance is called for here as if the data collection exercise is seen as leading to potential benefits then that may also lead to moral hazard and systematically biased measurement error.

The situation is different in relation to the assessments of children and families by teachers, doctors, health visitors, social workers and other front-line professionals and service providers. Here again there will be important ethical and administrative issues but measurement accuracy may be less biased than is the case for data from parents and other family members. Indeed these professionals may have access to very important information and judgements that were not available through the data collection of these national surveys.

To summarise, there is a very important and delicate relationship between the data collection element of a targeted policy implementation strategy for at risk individuals and families and the policy mechanisms triggered by that data collection. There are important issues about how the data is analysed, by whom and with what results that will impact not just on the success of the policy implementation but also on the quality of the data collected. It is likely that a good degree of experimentation and evolution of the system will be required and the information/policy relationship will be a key element of the design and evolution of this policy system.

3.3 Predictive measures

Table 3 provides summary statistics for all the variable and constructs used to assess risk, protection and development for the sample members and which are to be tested as signals of likely outcomes in the analysis described in subsequent chapters. Appendix Table 1 provides these by age and by the domains of measurement set out above.

In order to clarify the issue of the commensurability of measurement across domains and ages in the different datasets, Table 3 reports the number of constructs assessed at each age in each domain.

Table 3: Number of constructs, each dataset, by age and domain

	NCDS (1958 Cohort)				BCS (1970 Cohort)			
	0	7	11	16	0	5	10	16
Distal family factors	12	14	8	9	10	12	12	1
Proximal family processes	2	11	6	6	0	6	7	0
Mother characteristics								
Pre-school factors								
Neighbourhood context	0	0	1	0	0	2	0	0
School context	0	10	2	15	0	0	5	0
Cognitive development	0	9	10	10	0	7	4	0
Affective and behavioural development	0	15	19	16	0	3	14	12
Health	1	8	8	6	1	1	1	0

It can be seen, for example, that although the 1970 cohort has assessments of 12 different distal factors at age 5, at age 16 only 1 is assessed. Partly this is a restriction of the data collection, partly of what has been coded to date. It should be borne in mind when comparing the predictive power of distal factors at the two ages in this dataset. Other important points to note are that:

- Mother characteristics and features of pre-school contexts have not been included for the 1958 and 1970 Cohorts;
- The 1970 Cohort at age 16 only includes measures in the domains of affective and behavioural development, with the exception of one distal factor;
- Few measures of the neighbourhood context are assessed;
- Few health measures have been included, particularly for the 1970 cohort;
- Few school context measures have been included for the 1970 Cohort;
- In both Cohorts there are more measures of distal factors than of proximal processes.

3.4 Outcomes

In order to maximise the level of generality of the findings, analysis has been undertaken across the two datasets of a very wide range of outcomes. There are many important features of social exclusion and since particular measures may have differential predictive power for different outcomes, we have tested across a wide range of measures so as to provide as detailed an investigation of these relationship as possible and to limit the likelihood that results are only relevant to very specific single outcomes.

The outcomes were coded into a binary form, with positive outcomes denoting an outcome of adult deprivation. For each outcome sample members are defined either as positives (i.e. matching the social exclusion status denoted by the variable) or negative.

Appendix Table 2 sets out summary statistics for the set of outcomes or dependent variables for this study.

The outcome measures for the NCDS (1958 Cohort) are mainly assessed at age 42, some time after adolescence. In addition 7 outcomes have been coded at age 23 for that cohort, namely:

- Outcome 33, Age 23: Unskilled or semi-skilled
- Outcome 34, Age 23: Literacy or numeracy problem
- Outcome 35, Had child before 20
- Outcome 36, Had child before 23
- Outcome 37, Teen parent before 23
- Outcome 38, Depressed at 23
- Outcome 39, Low hourly wage at 23

The outcome measures for the 1970 Cohort are mainly assessed at age 30, although several refer to outcomes of adolescence, such as being a teen parent or experiencing high levels of time “Not in Education or Employment” (NEET) before age 18.

It must be noted that whereas some of the binary outcomes operate with a standard cardinal metric such that comparisons between the cohorts can be made, others are based on the relative position of individuals in the distribution in each dataset and so are not comparable. For example, the outcome of high level criminality (outcome 16) is defined as being found guilty in a court of law more than once. This is a relatively objective measure such that the increase in prevalence between the 1958 cohort (1%) and the 1970 cohort (5%) is interpretable as an increase in this outcome between the cohorts. There are two caveats: i) possible change in the reliability of self-reports over the period, ii) the measure for the 1958 Cohort is taken at age 42, whereas that for the 1970 Cohort is taken at age 30. These caveats make interpretation potentially hazardous but the raw data are certainly in line with reasonable assessment of the historical trend. That is not the case for the “NEET” variables (outcomes 29 and 30), for example. For these measures, the number of months during which each sample member was not in education, training or employment is assessed from self-reports. The binary variable indicates that the sample member experienced lengths of time in this state that were in the highest 10%. Therefore, by definition, the prevalence will be 10% and cannot change between cohorts. This definition is useful for the purposes of the analysis of this report but means that the measure does not provide information about the cross-period change. Note, too that the prevalence for outcome 29 in the BCS is 7%. That is because this proved to be a more amenable cut-off for the distribution of NEET status in those data. The cut-off has been selected to best characterise the data on NEET status and to identify a group for whom a negative outcome is most meaningfully recorded.

Multiple outcomes

We have also constructed indices of multiple deprivation. For the 1970 Cohort, we have constructed an index of 5 important outcomes, assessed at age 30, namely:

1. Mental health referral”
2. Social housing”
3. High levels of NEET before age 30
4. Overcrowding
5. Financial problem

Roughly 15% of the sample had 2 or more of these 5 outcomes (outcome 33), roughly 6% have 3 or more.

We have also constructed a measure indicating how many of the 32 outcomes each individual had experienced. We find that roughly 11% of the sample had experienced 9 or more of these 32 outcomes and we take this as a second high risk group (outcome 34). The correlation between these two indicators is fairly high at 0.61.

For the NCDS (1958 Cohort), we have constructed an index of multiple outcomes (outcome 40) based on use of 6 of the measures assessed at age 23 (all measures except low income which has a high rate of missingness). We use a cut-off of 3 of the 6 outcome measures, with 8.7% of the sample experiencing 3 or more of these 6 outcomes.

We also construct a second multiple outcomes (outcome 41) measure based on the age 42 data, indicating the 12% of the sample with more than 7 of the 32 outcomes assessed at that age.

4. Results of pathways analysis

4.1 Setting the scene: static predictions

The odds of adult outcomes:

Our starting point is a consideration of the static relationship between risk status at each age at which data is collected in childhood and the actual experience of adult social exclusion. The full set of results for all outcomes is presented in Appendix Table 3. A summary is shown in Table 4, below, for the multiple deprivation outcomes. Note that prediction is based on information at the indicated age only, not also accounting for information from prior ages.

Table 4. Odds of multiple adult deprivation given childhood risk categorisation

Outcome		age	Full information N	odds	t	Imputed N	odds	t
1958 Cohort								
40	>2 of 6 age 23 outcomes	0	4383	6.01	13.03	10123	4.22	16.32
		7	2682	13.37	14.57	8423	9.50	24.33
		11	3376	14.31	15.28	7647	11.07	24.76
		16	567	52.71	8.56	7424	11.75	25.30
41	>7 of 32 age 42 outcomes	0	6258	2.86	10.34	14673	2.53	14.57
		7	3823	5.65	13.59	12110	4.25	21.57
		11	4732	5.82	15.26	10828	4.89	22.75
		16	798	16.20	9.64	10560	5.32	23.82
1970 Cohort								
1	> 1 of 5 key outcomes	0	7974	4.02	18.76	10148	3.86	21.08
		5	5066	4.94	16.95	11112	4.36	24.42
		10	4565	6.25	18.81	11038	4.95	26.51
		16	4924	7.54	19.53	11112	4.43	24.67
2	> 8 of 32 outcomes	0	11412	2.88	15.76	16733	2.98	18.46
		5	6851	3.33	14.36	18779	4.18	26.35
		10	6126	7.93	23.83	18615	7.64	38.45
		16	6141	5.94	18.53	18725	5.72	32.83

It can be seen that the odds of experiencing the adult outcomes are very high for those who are defined as at risk at each age. Being at risk in childhood is very strongly associated with the likelihood of experiencing the indicated outcome. A substantial difference between the two cohort studies is that whereas for the 1958 cohort the odds for outcomes increases with the age at which risk is defined, for the 1970 cohort this effect is less marked. To some extent this is due to the weakness of the age 16 data collection in the 1986 survey of the 1970 Cohort, which coincided with a teachers' strike and also suffered other unusually great problems of attrition and missing data.

A second important distinction in Table 4 is the difference between results using the imputed data and those using only cases with full information. As discussed in the description of methods, above, identification is more accurate when only full information cases are used. Therefore, the odds ratios are much higher for this small sub-sample of the cohort than when cases are maintained by imputing data, this latter method maintaining cases in the sample for which little information is available, increasing the overall level of error in the forecast.

Nonetheless, a general finding across both cohorts and for both methods of treating missingness is that being classified as at risk at age 10/11 adds substantially to the already high risk of experiencing adult deprivation that follows at-risk status at birth.

For many outcomes (see also Appendix Table 3) in the 1958 cohort, risk status at age 7 is also a substantial augmentation to risk to that of birth. For example, in the imputed (lower bound) data, the odds of teen parenthood rise from 3.5 for risk status at birth to 6.6 for risk status at age 7, and are then fairly flat at 6.2, age 11 and 7.1, age 16. Equivalent age profiles are observed for many other outcomes. Age 5 risk in the 1970 Cohort, however, is for most outcomes, closer to birth risk than to age 10 risk in the strength of its negative implications for adult outcomes. This difference may be due to the lower general age gradient of these associations in the1970 than the 1958 Cohort. However, it is also a quite plausible inference that predictions made at age 7 are stronger than those at age 5 because children have spent longer in school leading to more useful teacher information in the formation of predictions and to the fact that the intervening two years are particularly important in lifecourse terms with a considerable degree of stabilisation of risk pathways. Important additional information about risk becomes available as children mature and signals provided by their own developmental pathways can be added to the prediction model

However, we should also recognise that substantial odds of adult deprivation are observed to follow from risk status at birth. In other words, for both cohorts, but particularly the 1970 Cohort, social address and family context are important markers of risk, development and opportunity. Thus, while remembering the great importance of developmental information, we can also recognise there is already important and actionable information about risk that can be drawn from knowledge of family circumstances at birth.

The persistence of risk

Before considering in detail the dynamic movements in and out of risk status, we first consider the number of periods at which individuals are so classified. As before the detailed results are in an Appendix Table (Appendix Table 4). For each outcome, in each dataset we report the distribution of the frequency of risk status. Examples, for the outcome of multiple deprivation, are shown graphically in Figure 2 for age 42 deprivation in the 1958 cohort, and in Figure 3 for the age 30 1970 Cohort. As well as showing the number of times children are classified as at risk, we also report in Appendix Table 4, how the likelihood of each outcome increases with the number of times in the childhood data at which individuals are said to be at risk. This likelihood gradient is reported in terms of both outcome probabilities and odds ratios. Example gradients are graphed in Figures 4 and 5 for the same multiple deprivation outcomes considered in Figures 2 and 3.

Figure 2: Number of sweeps "at risk," multiple deprivation at age 42, NCDS

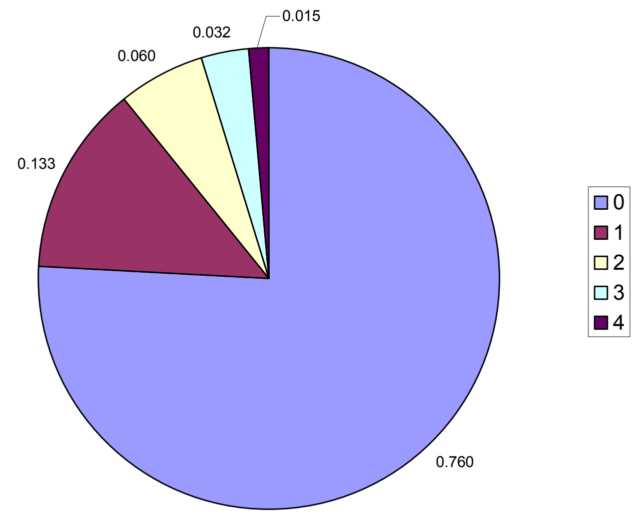


Figure 3: Number of sweeps " at risk," multiple deprivation age 30, BCS

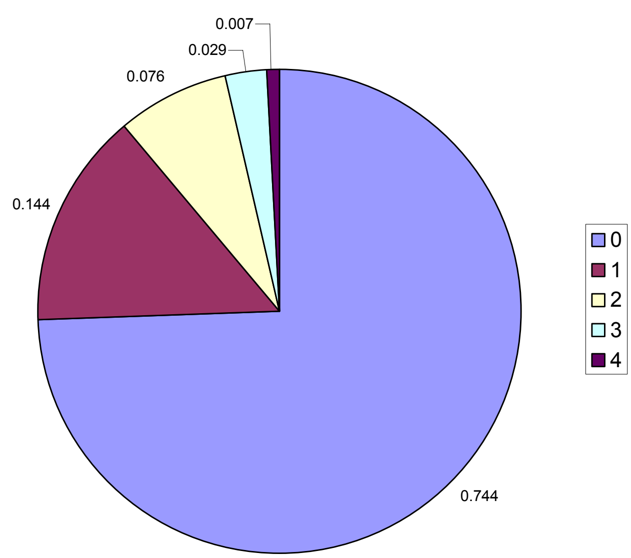


Figure 4: Probability of multiple deprivation age 42, NCDS, by no. of sweeps at risk

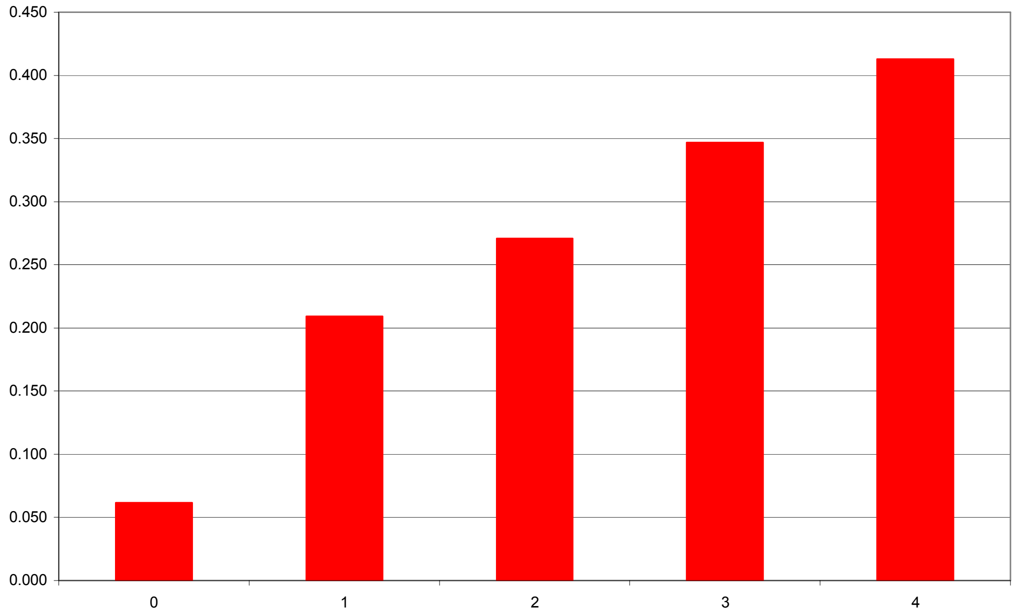
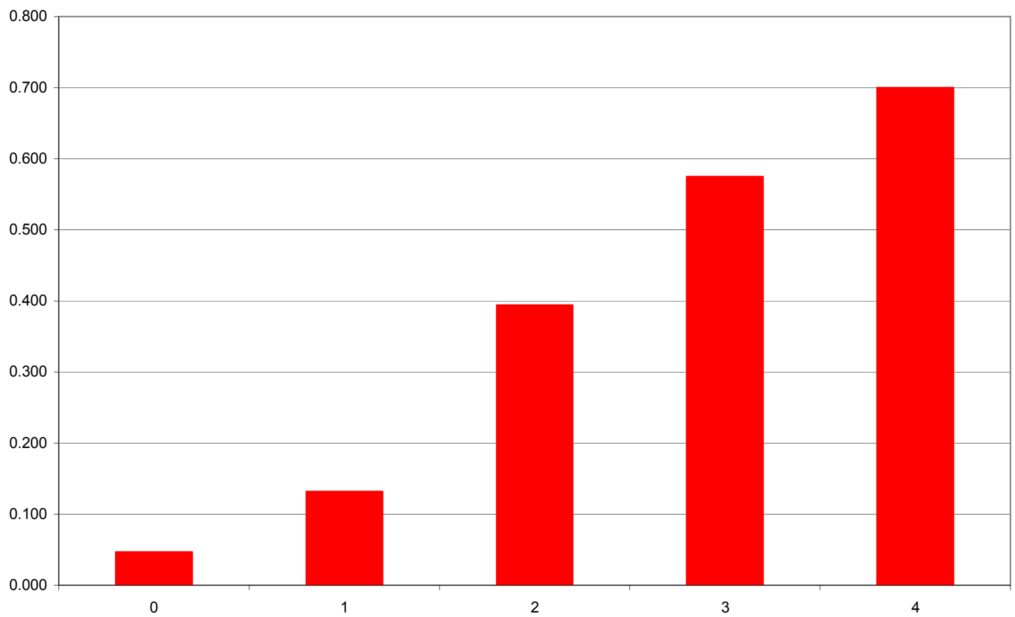


Figure 5: Probability of multiple deprivation age 30, BCS, by no. of sweeps at risk



The risk gradient is steep and fairly linear with respect to the number of episodes of risk. For example, the probability of experiencing the outcome of multiple deprivation at age 30 (Figure 5, 1970 Cohort) rises from 5% for those who were never classified as at risk in childhood to 39% for those who were at risk twice in the four ages at which data were collected and 70% for those who were classified at each episode.

Appendix Table 5 reports the persistence of risk for those who were classified as at risk in the first sweep or second sweep of each dataset. It can be seen that having been classified as at risk at birth, the likelihood of being reclassified as at risk is substantive but far from inevitable. In part this is due to the fact that different features of context and development were assessed at each age so that change in observed risk may be a matter of change in measurement focus rather than of a substantive real shift. However, these results also point to a substantial degree of churning, or movement in and out of risk. This is now assessed in more detail as we turn to the findings in the investigation of the dynamic patterns of risk.

4.2 Dynamic risk movements

Transition probabilities

Movements into and out of risk for each outcome and each dataset are shown in Appendix Table 6. Results for the multiple deprivation outcomes, using the imputed data, are reported below in Table 5.

Table 5. "At risk" transition probabilities, imputed data, 1958 and 1970 Cohorts

						Probability (excluding unknown cases)					
						Not at risk to:		At risk to:		To risk from:	
	Outcome	Age	Prior age	Mean risk prob	N	Not at risk	Risk	Not at risk	Risk	Not at risk	Risk
	1958 Cohort										
40	>2 of 6 age 23 outcomes	0		0.081	7941						
		7	0	0.083	8423	0.943	0.057	0.653	0.347	0.649	0.351
		11	0	0.082	7647	0.942	0.058	0.711	0.289	0.694	0.306
		11	7	0.082	7647	0.957	0.043	0.478	0.522	0.489	0.511
		16	0	0.084	7424	0.943	0.057	0.700	0.300	0.678	0.322
		16	7	0.084	7424	0.951	0.049	0.550	0.450	0.551	0.449
		16	11	0.084	7424	0.953	0.047	0.504	0.496	0.515	0.485
41	>7 of 32 age 42 outcomes	0		0.114	11403						
		7	0	0.116	12110	0.921	0.079	0.651	0.349	0.640	0.360
		11	0	0.117	10828	0.916	0.084	0.687	0.313	0.686	0.314
		11	7	0.117	10828	0.935	0.065	0.506	0.494	0.524	0.476
		16	0	0.118	10560	0.919	0.081	0.678	0.322	0.660	0.340
		16	7	0.118	10560	0.929	0.071	0.569	0.431	0.565	0.435
		16	11	0.118	10560	0.938	0.062	0.501	0.499	0.492	0.508
	1970 Cohort										
33	> 1 of 5 key outcomes	0		0.149	10148						
		5	0	0.151	11112	0.894	0.106	0.555	0.445	0.576	0.424
		10	0	0.151	11038	0.897	0.103	0.583	0.417	0.586	0.414
		10	5	0.151	11038	0.905	0.095	0.539	0.461	0.534	0.466
		16	0	0.151	11112	0.884	0.116	0.643	0.357	0.652	0.348
		16	5	0.151	11112	0.890	0.110	0.621	0.379	0.621	0.379

		16	10	0.151	11112	0.896	0.104	0.578	0.422	0.581	0.419
34	> 8 of 32 outcomes	0		0.117	16733						
		5	0	0.115	18779	0.914	0.086	0.615	0.385	0.628	0.372
		10	0	0.115	18615	0.902	0.098	0.711	0.289	0.719	0.281
		10	5	0.115	18615	0.921	0.079	0.613	0.387	0.609	0.391
		16	0	0.115	18725	0.899	0.101	0.749	0.251	0.753	0.247
		16	5	0.115	18725	0.912	0.088	0.673	0.327	0.673	0.327
		16	10	0.115	18725	0.921	0.079	0.605	0.395	0.607	0.393

Interpretation is as follows. Let us consider the first multiple deprivation outcome for the 1970 Cohort, namely the adult experience of more than 1 of the five key adult outcomes (outcome 33 from the full set). It can be seen that 14.9% of the 10,148 cases in the imputed sample experienced this. Of those not classified as at risk at birth, 10.6% were classified as at risk at age 5, 10.3% at age 10 and 11.6% at age 16. This figures can usefully be compared to the risk probabilities for those who were classified as at risk at age 0, of whom 44.5% were still at risk at age 5, 41.7% at age 10 and 35.7% at age 16.

Note that in Appendix Table 7 we report statistics that translate these figures into odds ratios for risk status at one age, given risk at the previous age. The relative odds ratios for the figures just reported are 6.8, 6.2 and 4.2

There are, therefore, two key findings:

1. There is considerable mobility. Many of those at risk at one age are not at risk at subsequent ages;
2. There is also important continuity in risk. The chances of being at risk at any age are much higher for those at risk at previous ages.

Overall, the findings here support the conclusion of a balanced perspective on risk continuity and discontinuity. Risk is not set in stone and immutable, neither is it random and unimportant for future chances.

Composition of the risk group

The final two columns of Table 5 show the composition of the group classified as at risk at each age, in terms of whether or not they were at risk at the prior age. Of those at risk at age 16, for example, for the same outcome as discussed above (no. 33), 34.8% had been classified as at risk at birth, but the remaining 65.2% were new cases[6]. Clearly any policy which highly emphasised risk at birth at the expense of continued monitoring and tracking with the potential for the adjustment and tailoring of resources in the light of new information would be inefficient.

[6] Note that the percentage of those at risk at age t who were also at risk at age t-1 represents also the true positive rate for the prediction from age t-1, i.e. the percentage of those who are at risk at age t who are identified by the prediction from the previous age. This P(TP) statistic is restated in Appendix Table 7, along with the associated false positive rate, P(FP).

The more specific figures are reported Appendix Table 6 for the full range of outcomes. The general lessons of the findings, however, are as discussed above and not restricted to this specific outcome and dataset.

Missing data

One other point to note is that the level of risk continuity is higher for the imputed data than for the full information data.

In the previous report (Feinstein and Sabates, *op. cit.*) on the static prediction from childhood risk status to adult deprivation, we described the prediction from the imputed data as a lower bound and the full information prediction as an upper bound. In that exercise the concern was with the accuracy in forecasting adult deprivation from childhood. When cases without full information were retained to form the prediction, this introduced a greater degree of error. The full information cases led to better prediction as they included greater information. However, they may overstate the degree of possible accuracy, depending on the method by which data might be collected in an equivalent policy exercise.

Here we find that there is less apparent continuity when only the full information cases are used. One explanation might be that when more information is allowed for, there is greater opportunity for new information to impact on the prediction and hence there is greater mobility. However, in part this may also result from the smaller sample sizes in the full information dataset which leads to greater variability and hence discontinuity in prediction.

The upper and lower bound interpretations are still useful as a guide to the likely range of continuity and discontinuity under different assumptions about missing data. However, for ease of presentation of so many statistics, we tend to report results from the imputed data because of concerns about the small sample sizes in the full information dataset and of non-random attrition. The broad generality of the findings is common across both methods, namely the finding of considerable discontinuity and mobility in risk, allied with substantial persistence for a core high risk group.

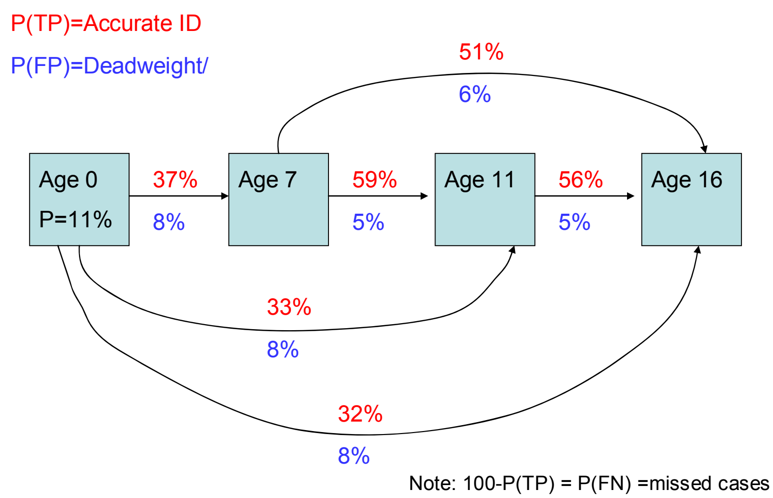
Using all constructs to assess period to period risk status continuity

An alternative way of considering risk continuity/discontinuity is to treat risk status at each age in childhood as an outcome to be predicted by information from previous ages and then to consider the rates of true and false positives in this prediction. This approach uses more information because instead of considering only risk status at the earlier age as a predictor of risk status at the later age, we use all information from all constructs at the previous age to predict risk status at the later age. This also leads to greater accuracy in prediction.

Results using this method are presented in Appendix Table 8 for all outcomes and also shown graphically in Figures 6-11 for a representative selection of outcomes, highlighting the use of the method for tracking risk between ages.

(Note that the numbers reported in Figures 6-11 are all contained in Appendix Table 8.)

Figure 6: Risk continuity; TP and FP for multiple adult deprivation age 23, NCDS



For example, Figure 6 shows the proportions of true and false positives (above and below the arrows, respectively) for transition from the situation of risk at one age to risk status at a later age. Of those classified as at risk at age 7, for the age 23 outcome of multiple deprivation in the 1958 Cohort, 37% would have been predicted to be so classified on the basis of their birth data (the true positive rate). Thus, 63% of those at risk at age 7 would not have been identified as at risk from the birth data, if an arbitrary cut-off was used to classify all cases at birth as either at risk or not at risk. Note, too that the false positive rate from 0-7 for this outcome is 8%, which is to say that 8% of those who were not at risk at age 7, would have been falsely predicted to be at risk from their birth data.

These figures demonstrate quite a high degree of discontinuity in risk between 0 and 7 and therefore inaccuracy in risk predictions made from birth. Figure 7-11 report the equivalent statistics for a range of other outcomes. It can be seen that in both datasets the true and false positive rates are much lower for the workless household outcome than for the outcome of mental health problem. This is mainly explained by the higher prevalence of the latter outcome. This means that more individuals are classified at risk, giving rise to higher numerators in the true and false positive probabilities

Figure 7: Risk continuity; TP and FP for workless household with children, age 42, NCDS

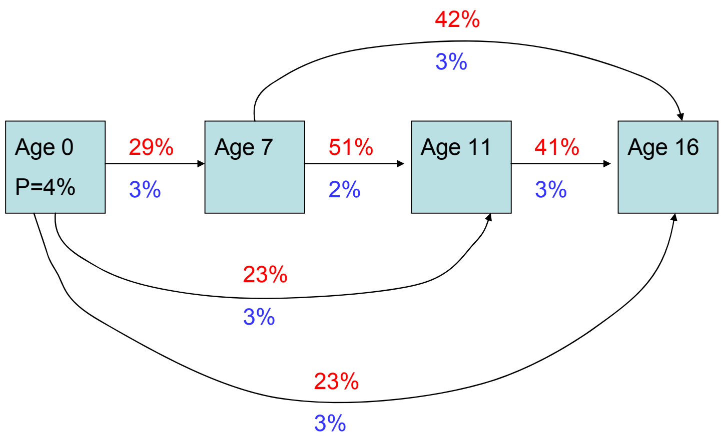


Figure 8: Risk continuity; TP and FP for mental health problem, age 42, NCDS

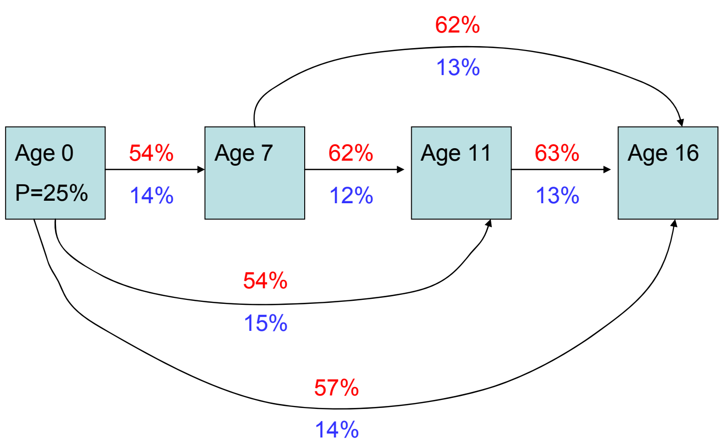


Figure 9: Risk continuity; TP and FP for multiple adult deprivation age 30, BCS

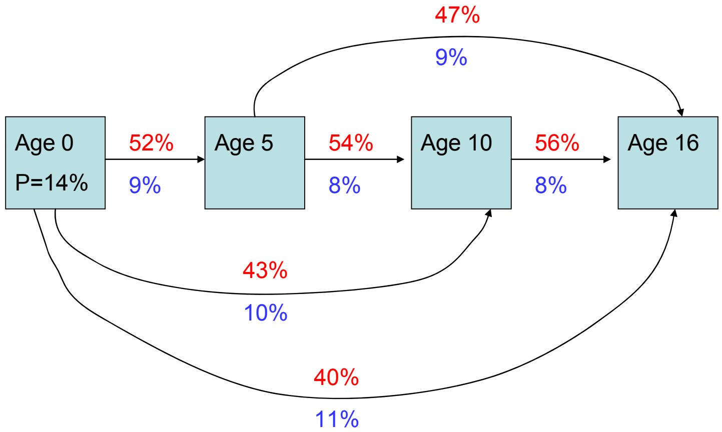


Figure 10: Risk continuity; TP and FP for for workless household with children, age 30, BCS

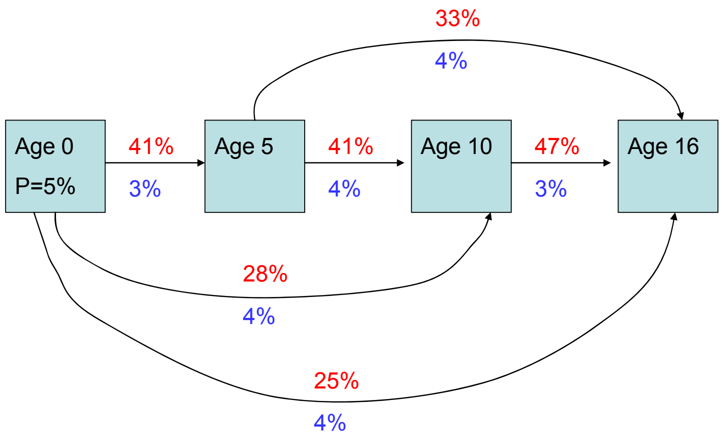
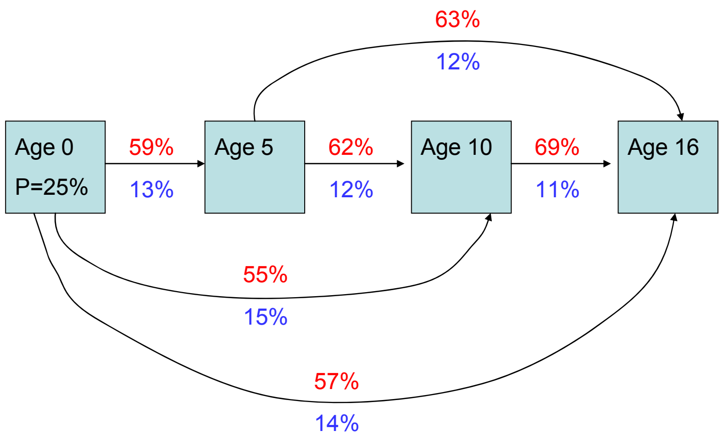


Figure 11: Risk continuity; TP and FP for mental health problem, age 30, BCS



Other important findings from Appendix Table 8 are:

1. There is much greater accuracy in risk forecasting in the 1958 Cohort from age 7 onwards than from birth onwards. For example, for the outcome of age 42 workless household, the true positive rate from 0-16 is 23%, that from 7-16 is 42%. There is clearly much greater continuity of risk from age 7 than from birth. This reflects the fact that the age 7 data includes wide-ranging assessments of the development of the sample member whereas the birth data only includes assessments of the context of the child. By age 7, many of the impacts of family risk are apparent (though not immutably) in the school attainments and behaviours of children and it becomes possible to better discern those children who are resilient to family risk or under-achieving despite less risky circumstances. These factors improve predictions considerably for the 1958 Cohort.

2. There is much higher continuity of risk from birth and so accuracy of forecast in the 1970 Cohort than in the 1958 Cohort. For example, for the multiple deprivation outcome, the true and false positive rates for the 1958 Cohort (from birth to age 23) are 37% and 8%; for the 1970 Cohort (from birth to age 30) the equivalent statistics are 52% and 9%. This is a big rise in the true positive rate without an equivalent rise in the false positive rate. These findings are mirrored across all outcomes.

3. There is not much greater stability or risk continuity from age 5 onwards in the 1970 Cohort than from birth onwards. Again, this represents a substantial difference between cohorts. Whereas (point 1 above) we find substantial increases in continuity of risk after age 7 for the 1958 Cohort, there is not a similar increase after age 5 in the 1970 Cohort. In part this follows from the much higher levels of risk continuity from birth for the later cohort so that the change is not explained by lower continuity from 5 onwards but by higher

continuity from birth onwards. It is feasible also that there is not a big increase in risk continuity in the 1970 Cohort from 5 onwards, as there is for the 1958 Cohort from 7 onwards, not only because of historical shifts in risk continuity per se, but also because of the age of assessment. Age 5 may be too early to assess developmental risk through child observation, particularly in terms of the accuracy of teacher assessments and written tests, which play a big role in risk assessment for the age 7 sweep. At age 5, tests and teacher assessments may be less accurate because i) children are still too young to sit reliable tests of academic achievement, ii) children are not yet so well known by the school and teachers, and, iii) children have not yet been strongly influenced by school environments. There may be important transitions from age 5 to age 7 that make age 7 a more appropriate age for accurate forecast of risk.

4. Risk continuity between sweeps at birth and age 16 is not much lower than that between birth and age 5/7. In some cases it is even higher. This reflects the impact of compounding risk. A fairly high proportion of those who exited risk status between birth and later ages, re-enter subsequently.

5. Calibrating programme effects on risk profiles

In the final section of this report, we consider the implications of the findings in Section 4 for the assessment of the relative costs and benefits of interventions targeted on those in need, at different ages through childhood.

The approach adopted makes a number of very strong simplifying assumptions which should be borne in mind in any presentation or inference from these findings. These are specified below and are useful in enabling us to clarify the trade-offs lost and gained through intervention at different ages, given what we have learnt through this analysis of the movements in and out of risk.

5.1 Method of calibration

Overview

We are interested in assessing the impact of interventions at each age at which risk status is observed in these datasets. The starting point for the calibration exercise is the recognition that intervention at any age in childhood will impact on the likelihood of adult deprivation and also on the likelihood of risk status at subsequent ages. The former channel leads to a direct benefit in terms of a reduction of the specific feature of deprivation. The latter, indirect benefit, is that fewer children will be at risk at the later age. This brings about a cost saving in that as well as reducing childhood deprivation and need, fewer children require subsequent intervention.

As children pass in and out of risk, intervention targeted only on those at risk at each age will miss those not at risk at any specific age but who nonetheless experience the adult outcome (the false negative cases). Thus, even with full effectiveness in targeted interventions there will still be incidence of adult deprivation. It should also be noted that some of the relevant underlying causal mechanisms may be driven by relative risks rather than absolute risk. For example, if jobs are rationed then increasing the skills of all and reducing the risk factors for the most disadvantaged children may lead to a change in the allocation of employment, but not necessarily in the overall incidence of unemployment. In the longer-term the level of employment is unlikely to be stable and may well increase with the overall skill level so the argument about the relative importance of absolute or distributional assets should not be over-stated. However, the point here is that removing the deprivation of the worst off may result in absolute benefits and equitable relative shifts but the precise impact on outcomes is not well understood and will also depend on the extent to which relativities are important in the formation of outcomes. This will also vary by outcome.

Similarly, there is no known intervention which will remove from society all features of deprivation and social exclusion. To some extent outcomes such as criminality and poor adult health will result either from social factors outside of the intervention range of specific programmes targeted on those in need and from personal and biological features that are not easily amenable to intervention. This is not to suggest that adult deprivation is inevitable or that pathways from early risk to adult deprivation are deterministic or inevitable. However, well-founded optimism about the capability of preventative intervention does have to be balanced with realism about its limits.

Another difficult issue for the calibration exercise is that of modelling plausible complexities of intervention heterogeneity with respect to differential effectiveness for different groups of the population or at different levels of universality or program breadth and roll-out.

Finally, in a more sophisticated model of intervention effectiveness, it would be appropriate to allow for differential effectiveness at different ages and the formal complementarities that follow from receipt at intervention at more than one age, which may add to overall effectiveness at a rate greater than the simple aggregation of the two treatment effects.

We abstract from these important issues in this analysis in order to address the main focus of the exercise, namely the implications of the risk dynamics observed earlier for the effective targeting of intervention and the assessment of the relative costs and effectiveness of intervention at different ages in childhood and adolescence.

Specific scenarios

Our approach is to assess the direct impact on adult outcomes and the indirect impact on programme cost under a range of intervention scenarios, namely:

1. Intervention for the risk group at each stage in childhood, impacting on a fixed proportion, f, of those at risk at each age. We consider i) the impact on the overall probability of adult deprivation for each outcome, and, ii) the impact on the proportions of those at risk at subsequent stages in childhood, allowing the effectiveness of intervention for the risk group, f, to take the values of f=0.1 (10%), 0.25, 0.5, and 1;

2. Intervention for the risk group with flexible effectiveness, rising or falling through childhood. We fix the average value of f over the four sweeps of intervention to be 1. In the model of increasing effectiveness, we set f to rise from 0.05 at birth, to 0.15 at age 5/7, 0.3 at 10/11 and 0.5 at 16. In the model of decreasing intervention effectiveness these values are reversed. Note, therefore, that the average level of effectiveness under either of these two models is equivalent to the constant rate model with f set at 0.25;

3. Intervention for the risk group at only one stage in childhood. In this scenario we consider what happens to the average probability of each form of adult deprivation and to subsequent risk probabilities if intervention only occurs at one age in childhood, thus assessing the relative value of equivalent investments at each different age.

A comparison of the direct benefit and indirect cost saving under this range of intervention scenarios enables an evaluation of the relative trade-off involved in setting different age profiles for intervention and of intervention at different ages.

Note that, as discussed further below, the parameter f can have a number of different interpretations under different assumptions as equivalently programme cost, universality or effectiveness. Assuming complete effectiveness then f also denotes universality, i.e. a programme that is delivered to 50% of the risk group will remove risk for 50% of that group. Alternatively, this can be thought of as a universal intervention for all of those at risk but with 50% effectiveness, removing risk for 50% of this group. An alternative interpretation is as cost, assuming linearity in effectiveness and universality, a programme with 100% effectiveness can be thought of as 10 times as expensive as one with 10% effectiveness. This latter interpretation is useful in providing a meaningful if arbitrary cost metric for the different scenarios. A more sophisticated model would allow universality, cost and effectiveness to enter the framework as three separate but related functions, and also build in diminishing returns to universality, differential effectiveness at different ages and formal individual complementarities of multiple intervention. We leave that to later work.

Modelling strategy

The impact on the overall probability of adult outcomes can be simply modelled using the notion of risk profiles. The risk status indicator for each child at each of the four ages of data collection can take a value either or 0 (not at risk) or 1 (at risk.) Therefore, there are 16 (2^4) possible profiles of risk, varying from a profile such as 0000 for a child never defined as at risk to 1111, for a child classified as at risk at each sweep.

The impact of intervention will very depending on the risk profile. For example, an individual with a profile 0001 will only benefit from intervention at the last sweep, age 16. An individual with a profile 0101 will benefit from two episodes of intervention. The overall impact on adult outcomes of different scenarios of intervention through childhood will differ for each risk profile but the overall post-intervention probability of the adult outcome will be the sum across all risk profiles of the product of the probability of the risk profile and the probability of the outcome for individuals with the risk profile.

For example, the contribution to the overall probability of the outcome of those with profile 0100, i.e. only at risk at the second sweep is given by:

$$P\ (Y_{0100}) = P(R_0=0, R_1=1, R_2=0, R_3=0) * P(Y=1 \mid R_0=0, R_1=1, R_2=0, R_3=0)$$

The overall probability is given by:

$$P\ (Y) = \sum\sum\sum\sum P(R_0=r_0, R_1=r_1, R_2=r_2, R_3=r_3) * P(Y=1 \mid R_0, R_1, R_2, R_3)$$

summing across risk status at all four ages, i.e. for all risk profiles, with r_t taking the values of 0 or 1 at each age. We denote the first sweep as t=0 to highlight the fact that this data collection occurs at birth.

Introducing intervention effectiveness into this formula, we state that under intervention the probability of the outcome for those receiving the intervention is reduced by the factor $(1-f_t)$ where denotes programme effectiveness at age t. A universally effective intervention with a vale of $f_t=1$ will reduce the probability of the adult outcome to 0 for the individuals who receive the intervention by virtue of being classified as at risk at age t.

Thus, with interventions introduced, the overall probability of the outcome for the specific risk profile 0100 becomes:

$P(Y_{0100}) = P(R_0=0, R_1=1, R_2=0, R_3=0) * P(Y=1 \mid R_0=0, R_1=1, R_2=0, R_3=0) * (1-f_1)$

If programme effectiveness at time 1, the second sweep, is complete then the contribution to the overall probability of the outcome for those with this profile will be zero.

Summing, as before, we obtain:

$$P(Y) = \sum\sum\sum\sum P(R_0, R_1, R_2, R_3) * P(Y \mid R_0, R_1, R_2, R_3) * \prod (1- f_t * r_t)$$

Here, for each profile the contribution to the overall probability of the outcome is reduced by the factor $(1- f_t)$ for each episode at which those with the profile are at risk and receive the intervention.

Equivalent formulae are derived for the indirect benefit of intervention in terms of reducing the probability of risk status at later ages, i.e. for $P(R_3=1)$, $P(R_2=1)$ and $P(R_1=1)$ given the set of risk profiles prior to each of these ages. Thus, for example, for R_1, i.e. risk at age 5/7, the only relevant prior profiles are 01 and 11 as:

$P(R_1=1) = P(R_1=1 \mid R_0=0) * P(R_0=0) + P(R_1=1 \mid R_0=1) * P(R_0=1)$

Intervention at birth will only occur for the profile 11. Thus, the post-intervention probability of risk status at sweep 2 is given by:

$P(R_1=1) = P(R_1=1 \mid R_0=0) * P(R_0=0) + P(R_1=1 \mid R_0=1) * P(R_0=1) * (1-f_0)$

If intervention at birth is fully effective in removing the risk of later risk status for those at risk at birth, $f_0=1$, there will still be individuals classified as at risk at the second sweep, with a probability given by $P(R_1=1 \mid R_0=0) * P(R_0=0)$, i.e. the likelihood of new entrants to risk.

Equivalent formulae are derived for risk status at later ages. Results using these formulae are presented in the next section.

It should be remembered that key assumptions are:

1) That the identification of risk in data collection to inform an equivalent "real world" exercise would be comparable to that observed here. In fact, these data may underestimate the potential of administrative data and other professional and practitioner observations to accurately identify cases of need. However, the impact of a broader and more localised data collection system on the age profiles and hence on the calibration results is hard to predict a priori;

2) That the effectiveness of policy is similar to that modelled here, with all the linearity and homogeneity assumptions implicit in this simple framework.

3) That the impact of risk factors is due to absolute effects of deprivation not to relative aspects.

5.2 Calibration results

The full set of results for all outcomes for the direct effect of intervention under the different scenarios is reported in Appendix Table 9. The results for the indirect benefits are reported in Appendix Table 10 and 11. Combined results for effects on the overall benefit to cost ratio are reported in Appendix 12.

Direct effects on adult outcomes

Results for the multiple deprivation outcomes are reported in Table 6.

Table 6: Direct impact of intervention on the probability of adult outcomes

	Probability of adult outcome										
	No inter-vent-ion	**Constant f**				**Flexible f, average=50%**		**One age only, f=100%**			
Outcome		**f=1**	**f=0.5**	**f=.25**	**f=0.1**	↑	↓	**Age 0**	**Age 5/7**	**Age 10/11**	**Age 16**
NCDS											
Imputed dataset											
40: >2 of 6 age 23 outcomes	0.075	0.027	0.040	0.054	0.066	0.037	0.041	0.055	0.047	0.048	0.045
41: >7 of 32 age 42 outcomes	0.108	0.047	0.067	0.084	0.097	0.063	0.066	0.083	0.080	0.076	0.076
BCS											
Imputed dataset											
33: > 1 of 5 key outcomes	0.149	0.046	0.076	0.104	0.129	0.071	0.075	0.099	0.093	0.092	0.094
34: > 8 of 32 outcomes	0.117	0.040	0.064	0.085	0.103	0.058	0.067	0.088	0.081	0.070	0.076
Full information dataset											
33: > 1 of 5 key outcomes	0.106	0.035	0.054	0.075	0.092	0.048	0.058	0.076	0.069	0.065	0.064
34: > 8 of 32 outcomes	0.099	0.034	0.055	0.073	0.087	0.049	0.059	0.075	0.070	0.059	0.067

Note that if the cost of intervention relative to effectiveness is constant across ages then one can see the effectiveness parameter is an indicator of cost. An intervention that achieves effectiveness given by f=1 at each age has an overall cost according to this indicator of 4. The cost of a programme provided only at one age but with effectiveness of 100% is 1 by this metric. The cost of this particular multiple episode intervention scenario is 4 times that of this one-period design. Similarly, the cost of a 4 period design with f=0.25 at each age can be thought of as equivalent to the single age design. The difference in benefit can be observed in Table 6. There are a number of interesting findings. It should be remembered that so far we have not taken account of the indirect benefits.

1. There are decreasing returns to expenditure, f. A programme at each age at 10% effectiveness would in this model reduce the likelihood of age 23 multiple deprivation for the 1958 Cohort from 7.5% to 6.6%, a reduction of 0.9 percentage points. The reduction from a programme with f=0.25 brings about a reduction of 2.1 points, at f=0.5 the reduction is 3.5 points. As f increases, the rate of decline of the adult outcomes falls.

2. A programme of expenditure at one age only at 100% cost/effectiveness can be compared to an equivalent programme of 25% cost at each of the four ages. It can be seen that in terms of the direct benefit of reduced adult deprivation, the fixed age intervention is more effective. Thus 100% intervention only at age 11 reduces the overall probability of outcome 33 in the 1970 Cohort to 9.4% as compared to the benefit of a 25% intervention at each age which is a reduction to 10.4%.

3. For all outcomes, intervention only at birth is less effective than intervention only at any subsequent age. On the whole the direct outcome benefits are greatest for intervention only at 10/11 or 16. This results from the greater accuracy of later targeting.

4. The direct benefit in terms of reduced adult deprivation of intervention that increases in cost/universality through childhood is greater than that of a programme of higher early expenditure that falls through childhood. Thus, for example, the overall probability of outcome 33 in the 1970 Cohort is reduced from 14.9% without intervention to 7.1% under a programme with rising expenditure/effectiveness through childhood. This compares to 7.5% with a programme that decreases through childhood.

Indirect effects on later childhood risk status – cost reduction

Turning now to the important indirect effects of intervention, the full set of results in relation to impacts on later risk status probabilities are reported in Appendix Table 10 and specific results for multiple deprivation presented below in Table 7. (We present only a sub-set of results here for ease of presentation but the general results discussed are similar across outcomes.)

Table 7: Indirect impact of intervention on the probability of later childhood risk status

	NCDS: >2 of 6 age 23 outcomes			BCS: > 1 of 5 key outcomes		
	P(R7=1)	P(R11=1)	P(R16=1)	P(R7=1)	P(R11=1)	P(R16=1)

No intervention	0.077	0.072	0.078	0.156	0.150	0.152
f=1	0.047	0.027	0.025	0.090	0.053	0.055
f=0.5	0.062	0.045	0.042	0.123	0.092	0.087
f=0.25	0.070	0.058	0.057	0.140	0.119	0.114
f=0.1	0.074	0.066	0.069	0.149	0.137	0.136
Flexible f: Rising	0.074	0.059	0.049	0.149	0.123	0.100
Flexible f: Falling	0.047	0.035	0.037	0.090	0.067	0.075
Age 0 only	0.047	0.047	0.053	0.090	0.088	0.099
Age 7 only	0.077	0.035	0.042	0.156	0.078	0.094
Age 11 only	0.077	0.072	0.043	0.156	0.150	0.089

Table 7 highlights a very different set of findings to that for the direct benefits. In particular:

1. A programme profile of high initial expenditure, falling to lower levels through childhood has a bigger impact than the schedule of a rising profile. This is because early intervention reduces later risk and these benefits are compounded through later investment. The early intervention brings greater opportunity for complementarity.

2. Equivalently, it is also the case for investment at one age only that the lowest average level of risk follows from the earliest investment.

Resulting cost savings

The indirect benefit of reduced later childhood risk brings about cost savings in that lower levels of intervention are required at later ages. Thus, for example, as the first row of Table 7 indicates, an intervention targeted on those at risk of age 23 adult deprivation, with 100% universality for the age 7 risk group must be targeted on 7.7% of the overall population if there has been no prior intervention. However, an intervention schedule which reduces risk by 50% for those at risk at birth, leads to a reduction in the risk group at age 7 to 6.2%. This must be factored in alongside the direct benefit of that programme schedule.

Therefore, in Appendix Table 11 and Table 8, we present results on the cost saving that follows from this reduced need for subsequent intervention.

Table 8: Reduction in cost of intervention schedule, resulting from reduced risk later in childhood, as % of total

	Constant f				**Flexible f, average=50%**		**One age only, f=100%**			
Outcome	**f=1**	**f=0.5**	**f=.25**	**f=0.1**	**↑**	**↓**	**Age 0**	**Age 5/7**	**Age 10/11**	**Age 16**
NCDS										
Imputed dataset										
40: >2 of 6 age 23 outcomes	3.21	1.95	1.08	0.46	1.02	0.90	0.00	0.00	0.00	0.00
41: >7 of 32 age 42 outcomes	4.24	2.56	1.40	0.59	1.36	1.14	0.00	0.00	0.00	0.00
BCS										

Imputed dataset										
33: > 1 of 5 key outcomes	6.51	3.91	2.14	0.90	1.86	1.93	0.00	0.00	0.00	0.00
34: > 8 of 32 outcomes	4.42	2.58	1.40	0.58	1.30	1.24	0.00	0.00	0.00	0.00
Full information dataset										
33: > 1 of 5 key outcomes	3.91	2.25	1.21	0.50	1.06	1.14	0.00	0.00	0.00	0.00
34: > 8 of 32 outcomes	3.15	1.82	0.98	0.41	0.83	0.91	0.00	0.00	0.00	0.00

We note three findings in particular:

1. The relative benefit from a wider programme (higher f) is increasing with respect to f, counter to the implication for model of the direct benefit. Thus, the % reduction in cost is higher for a programme with a high reach than for one impacting a lower proportion of the eligible population.

2. The reduction in cost is greater here for a programme with a rate of intervention that increases through childhood, rather than for one which decreases through childhood. This is because although the declining schedule of a high early intervention programme leads to a greater reduction in later need than a schedule with high late intervention, the benefits in terms of reduced later cost are factored down as fewer individuals will be targeted by a late programme in the falling intervention schedule. Thus, there is a trade-off between two factors and for these schedules modelled here, the fact that fewer receive the late intervention in a declining schedule outweighs the greater reduction in the probability of need. The precise relative value of declining and rising schedules will depend on the specific weights given to intervention at each age and for some schedules a declining schedule would lead to a greater cost saving than a rising schedule. However, the key implication to emphasise here is that the high early intervention schedule does not necessarily lead to a cost saving greater than that for a rising schedule.

3. There are no cost savings resulting from the indirect benefit of reduction in later risk, for a schedule with only impacts at one age. This follows automatically from the one-period approach as there is no subsequent intervention in such a design.

Integration of the direct and indirect benefits of intervention profiles

In this section we report the cost/benefit ratio obtained for each intervention scenario. We measure cost in terms of the parameter f, adjusted for later sweeps in childhood for any cost savings resulting from earlier intervention that reduces the probability of later risk status (as in Table 8, Appendix Table 11). Benefit is assessed in terms of the metric of reduced probability of the adult outcome.

Results are presented in Appendix Table 12 and for multiple deprivation in Table 9.

Table 9: Benefit/ cost ratio factoring in both direct and indirect benefits of intervention

	Constant f				Flexible f, average=50%		One age only, f=100%			
Outcome	**f=1**	**f=0.5**	**f=.25**	**f=0.1**	**↑**	**↓**	**Age 0**	**Age 5/7**	**Age 10/11**	**Age 16**
NCDS										
Imputed dataset										
40: >2 of 6 age 23 outcomes	0.007	0.021	0.054	0.165	0.052	0.056	0.055	0.047	0.048	0.045
41: >7 of 32 age 42 outcomes	0.012	0.034	0.085	0.245	0.083	0.084	0.083	0.080	0.076	0.076
BCS										
Imputed dataset										
33: > 1 of 5 key outcomes	0.012	0.039	0.107	0.325	0.104	0.106	0.099	0.093	0.092	0.094
34: > 8 of 32 outcomes	0.011	0.033	0.086	0.258	0.083	0.089	0.088	0.081	0.070	0.076
Full information dataset										
33: > 1 of 5 key outcomes	0.009	0.028	0.076	0.231	0.072	0.078	0.076	0.069	0.065	0.064
34: > 8 of 32 outcomes	0.009	0.028	0.074	0.219	0.071	0.076	0.075	0.070	0.059	0.067

The results in Table 8 suggest that on the while, marginally the most effective intervention schedule in terms of the indirect benefit derived from reducing subsequent risk and hence the cost of subsequent intervention, is a schedule with high initial intervention, reducing in cost/universality through childhood. However, this is only slightly more beneficial than a schedule with a flat profile. For the outcome of multiple deprivation in terms of the 5 key outcomes for the 1970 Cohort, a flexible intervention scheme on a declining schedule with an average cost of f=0.25 over the four periods, has a benefit : cost ratio of 7.8%. This compares with the ratio from a flat schedule with the same average cost of 7.6% points and a ratio from a rising schedule of 7.2%.

We also note that if there is to be intervention at only one age, the most effective such age is birth.

Finally, we note that the overall benefit/cost ratio declines with the level of effectiveness, f. This follows from the diminishing direct returns to f, observed in the consideration of the effects of the different intervention schedules on the probability of the outcomes of adult deprivation. This is a direct consequence of the specification of intervention effectiveness in this simple framework. We would recommend that future work consider more advanced models than can consider the results under a range of different specifications that loosen the strong simplifying restrictions imposed in this basic framework. It should be remembered, too, that in this analysis, we have not allowed for greater effectiveness of early intervention. This, too, should be modelled in subsequent calibration.

To conclude, we find that there are three conflicting factors which influence the relative benefit/cost ratio of different age profiles for intervention in the different scenarios modelled:

i. Later intervention is more effective because it can be targeted more easily;
ii. Earlier intervention is less costly because it reduces the need for later intervention;
iii. A schedule with high late intervention is less costly because it obtains greater benefit from the reduction in cost of late intervention, than a schedule with high early intervention.

Overall, we find that a slightly declining schedule of intervention is most effective but that a flat schedule is not much less cost-effective and in the case of many outcomes more so.