Lifecourse Economic Evaluation: A Microsimulation Approach

Ieva Skarda[a], Miqdad Asaria[b] and Richard Cookson[a]

May 15, 2019

Abstract

We present LifeSim – a new framework for lifecourse economic evaluation, which provides detailed information about long-term costs, benefits and inequality impacts in terms of health, consumption and wellbeing over the lifecourse. We use life-stage-specific networks of equations to simulate the causal pathways linking early life circumstances and skills formation to diverse later life outcomes. Our equations from age 0 to 14 are primarily based on longitudinal survey data from the Millennium Cohort Study. Later life equations are parameterised using causal effect estimates from quasi-experimental studies combined with target data from surveys and administrative records, with outcomes to age 46 validated using the 1970 British Cohort Study. We illustrate the framework by evaluating a training programme for parents of young children at risk of conduct disorder. We trace how multiple disadvantages cluster and compound over time to generate heterogeneity in long-run benefits and costs, allowing us both to pinpoint which subgroups benefit most and to simulate distributional impacts on inequality of opportunity for lifetime health and wellbeing within the general population.

Keywords: Simulation modelling, Cost Benefit, Health, Human capital, Skill, Inequality, QALY, Quality of life, Well being

JEL Classification: I14, I24, I31, J24, C53, C63, D61, D63

---

[a] Centre for Health Economics, University of York, Heslington, York, YO10 5DD, UK

[b] LSE Health, Cowdray House, London School of Economics and Political Science, Houghton Street, London WC2A 2AE, UK

Correspondence: Ieva Skarda, Centre for Health Economics, University of York, Heslington, York, YO10 5DD, UK; email: ieva.skarda@york.ac.uk; tel: 01904 321475

This is independent research supported by the National Institute for Health Research (SRF-2013-06-015). The views expressed in this publication are those of the authors and not necessarily those of the NHS, the National Institute for Health Research or the Department of Health.

1 Introduction

Many public policies have important but poorly understood long-run consequences for individual health and wellbeing, public cost and social inequality. A paradigmatic example is early years policy to improve child development (Doyle et al., 2009; Heckman, 2006). Trials and quasi-experiments can provide useful information about short-run policy impacts, such as effects on children's cognitive and socio-behavioural skills (Scott et al., 2014; Hutchings et al., 2007; Gardner et al., 2006). However, simulation modelling is required to extrapolate these impacts into adulthood and quantify the long-run consequences of ultimate interest to policy-makers (Heckman et al., 2010; García et al., 2017; Garces et al., 2002; Schweinhart et al., 2005). As well as improving understanding of long-run policy impacts for the average citizen, lifecourse microsimulation can also improve understanding of policy impacts on inequality. Distributional impact analysis often only provides a snapshot of inequality at a point in time, which can potentially be misleading if the underlying policy concern relates to inequalities in people's lifetime opportunities and outcomes (Hills, 2017).

Our aims in this paper are two-fold:

(i) to introduce LifeSim – a prototype lifecourse microsimulation model of an English birth cohort, capable of simulating both baseline levels of, and policy effects on, diverse health and non-health outcomes and public costs over the lifecourse, and;

(ii) to show how LifeSim can be used to conduct lifecourse economic evaluation of early years policy, using the illustrative example of a training programme for parents of children aged 5 showing signs of antisocial behaviour (Gardner et al., 2017).

We see this a modest contribution towards the larger international research endeavour of building detailed, credible and user-friendly lifecourse microsimulation models that policy-makers can use routinely to gain a better understanding of the long-term costs, benefits and inequality impacts of public policies (Layard et al., 2014; Frijters et al., 2017).

The standard approach to quantifying long-run impacts in applied cost-benefit analysis (CBA) is what we call the "separate impact modelling" or "multiplier" approach (Lee et al., 2012). In this approach, different long-run impacts are measured and monetised using separate production functions estimated using information about population average relationships. In practice, this often boils down to the application of long-run “multipliers” to short-run effects (Paull and Xu, 2017). For instance, preventing a case of child conduct disorder might be converted into an average long-run number of crimes prevented and then monetised using a cost to society per crime; an increase in cognitive skills might be converted into an average increase in educational attainment and then converted into an average increase in lifetime earnings; and an increase in educational attainment might also be converted into an increase in life expectancy and then monetised using a value per life year gained. A key limitation of this approach is that the separate production functions take no account of how different impacts (e.g. on education, crime, earnings, health) cluster within the same individual and interact with one another as they jointly evolve over the lifecourse. The “multiplier” approach can lead to over-estimation of benefits due to double counting, or under-estimation through failing to allow for the clustering of disadvantages, and it is not easy to assess either the direction of magnitude of bias. Accurate quantification of long-run consequences therefore requires individual-level microsimulation of life histories, which jointly models how different consequences evolve over the lifecourse (Zucchelli et al., 2012).

Our study contributes to the literature in three ways. First, we introduce a novel lifecourse microsimulation framework for economic evaluation, that provides detailed individual-level information about multiple long-run health and non-health effects and costs based on careful modelling of causal pathways. Second, we illustrate a novel approach to comparing and valuing health and non-health effects in terms of years of good life gained. Conventional monetary valuation based on willingness to pay can be criticised for bias in favour of the rich, since it ignores variation between people in the marginal value of income. We use an alternative summary wellbeing metric based on the quality adjusted life year (QALY) concept from health economics. Our measure extends the health QALY by adjusting quality of life not only for health quality but also for income (Cookson et al., 2016). Third, we illustrate a novel approach to evaluating policy impacts on inequality, by using our detailed life history simulations to construct summary measures of policy impact on equality of opportunity for lifetime wellbeing.

LifeSim provides information on baseline life histories as well as the marginal effects of policy change. This allows us to provide a full, integrated picture of inequality levels under different policy scenarios. An alternative approach would be to focus on modelling marginal effects, and then describe baseline life histories separately by piecing together existing survey and mortality data with as few modelling assumptions as possible (Frijters et al., 2017).

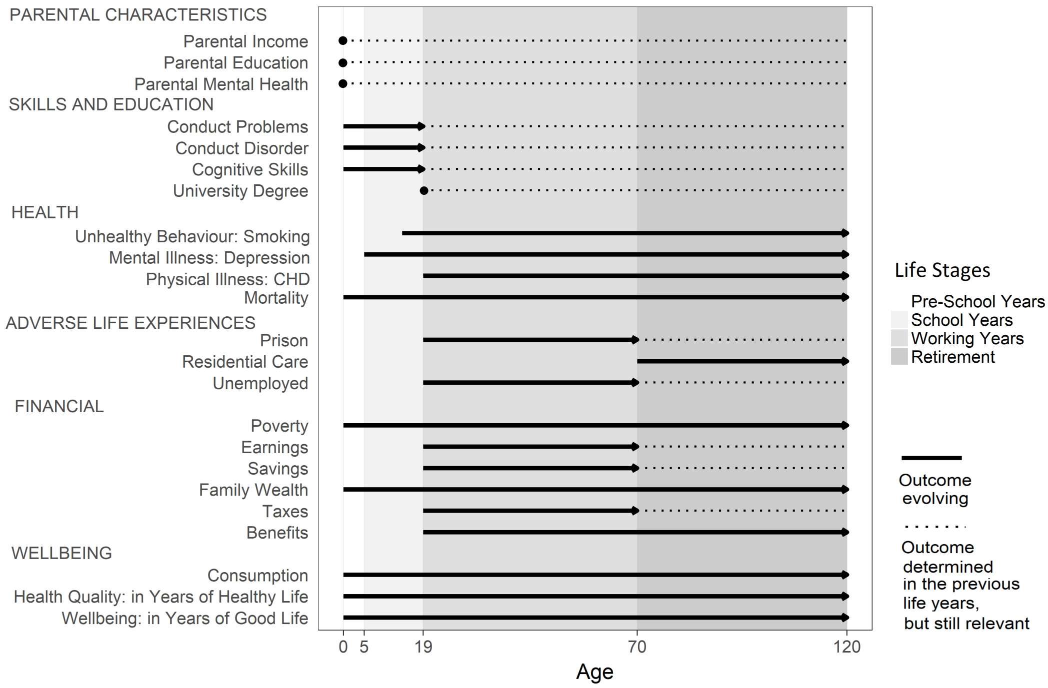
LifeSim simulates life histories for a whole-population cohort of almost 100,000 English children born in 2000-1. Outcomes up to age 14 are primarily based on data from the Millennium Cohort Study (MCS). We then model the year-by-year evolution of lifecourse outcomes after age 14 using life-stage-specific equations, drawing on inter-disciplinary theory, quasi-experimental evidence and various sources of target data for long-run consequences. We compare our model predictions with data from the 1970 Birth Cohort Study up to age 46 as a simple validation check. We then use LifeSim to assess the long-term consequences of a parent training programme designed to improve socio-behavioural skills in children age 5 screened as having high-risk of conduct disorder. We do this to illustrate how our approach can add useful new information for policymakers, beyond the information that can be provided by conventional methods of cost-benefit analysis.

2 Methods

Our microsimulation modelling framework specifies an integrated network of life-course-specific causal pathways linking diverse individual-level life outcomes of interest to policy makers. Central to the framework is basic human capital formation in childhood and how this influences educational attainment, earnings, physical illness, mental illness, mortality and other outcomes with important impacts on individual wellbeing and public cost (see figure 1). The structure of causal pathways changes as individuals progress through four key life stages: pre-school years (age 0-4), school years (age 5-18), working years (age 19-69) and retirement years (age 70+). As an example, we illustrate the working age pathways in figure 2; other pathways are shown in the Supplementary Appendix A.

Our framework distinguishes between two types of basic human capital: cognitive and social skills. Cognitive skills are the basic information-processing skills that the brain uses to think, learn, remember, reason, and pay attention. Social skills enable people to communicate, learn, ask for help, get needs met in appropriate ways, get along with others, protect themselves, and in general, interact harmoniously with others. Our prototype model focuses on one specific type of social skills – good conduct – though the framework can in principle be extended to include multiple social skills. Child conduct is related to self-control and regulation, which have been shown to matter in many aspects of life, including wellbeing, income, employment, crime and health outcomes (Joshi et al., 2016).

Figure 1: Overview of Key Outcomes over the Lifecourse

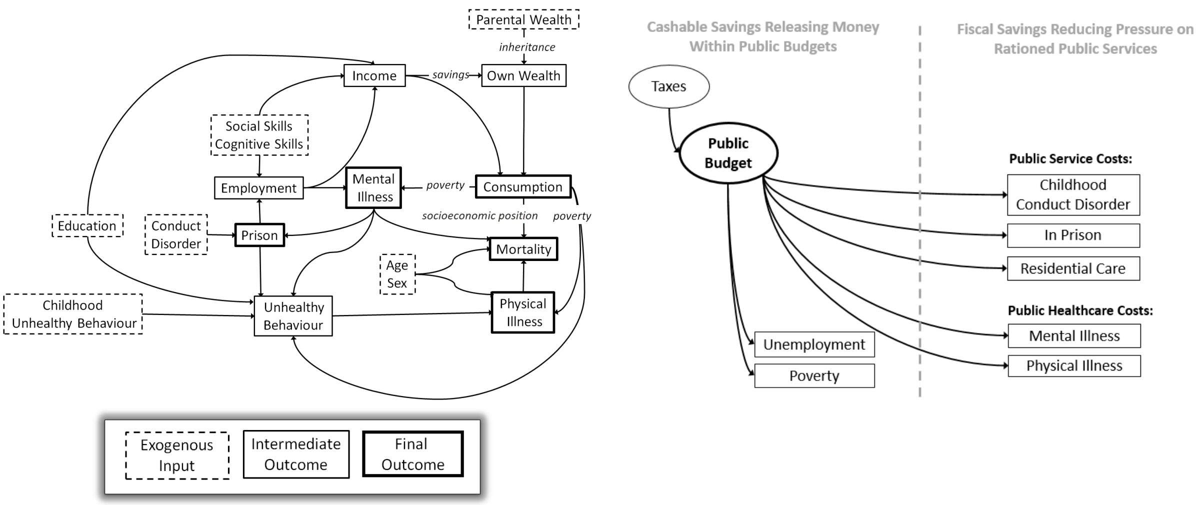


Note: the modelled outcomes throughout the key life stages: preschool years (up to age 4), school years (ages 5-18), working years (ages 19-69), and retirement (age 70+). The solid arrow represents the years when a variable is evolving. The dotted line represents the years when the variable has stopped evolving, but the level determined in past remains relevant, e.g. university degree is determined at 19, but stays relevant throughout the entire life.

Figure 2: Causal Pathways and Financial Flows

(a) *example of causal pathways: "working years"*

(b) *public budget*



We implement our framework by simulating a cohort of 100,000 English children born in 2000-

2001. We use data from the Millennium Cohort Study (MCS) to describe their initial characteristics and family circumstances – including parental wealth, education and mental health. We also use MCS data to describe outcomes up to age 14, supplemented where necessary by modelling rare childhood outcomes such as mortality using external data sources. We then model the causal pathways leading to later life outcomes using equations parameterised using quasi-experimental evidence and calibrated against various external sources of target data.

2.1 Childhood Survey Data

We use random draws of individuals from the MCS to represent the initial conditions of the simulated cohort, as well as the child's cognitive skills and conduct problems. The definitions and summary statistics of the skills measures and initial conditions are summarised in table 1.

We measure conduct problems during childhood using the Strengths and Difficulties Questionnaire (SDQ) conduct problems subscale score. This score ranges from 0-10, with a higher score representing more conduct problems. As suggested in the SDQ official website[1], we use the parent-reported SDQ conduct score cut-off value of 4 or above as a simple screening indicator of children at risk of developing a conduct disorder, to select which parents are offered parent training. However, we model the actual probability of developing conduct disorder using a more sophisticated predictive algorithm based on a combination of SDQ conduct score and a further parent-reported "behavioural impact" score, which provides a specific probability of conduct disorder based on a classification as either "possible" or "probable" (Goodman et al., 2003, 2000).

[1]The documentation can be downloaded from `http://www.sdqinfo.org/py/sdqinfo/b3.py?language=Englishqz(UK)`

Table 1: Summary of Skills Measures and Family Conditions in the Childhood Survey Dataset

Initial Condition	Data Used in Modelling	Source	N	Mean	SD	Min	Max
Social Skills							
Conduct Problems up to Age 4	SDQ conduct problems score at age 3	MCS2	4023	2.64	1.94	0	10
Conduct at Ages 5-6	SDQ conduct problems score at age 5	MCS3	4023	1.32	1.38	0	8
Conduct at Ages 7-10	SDQ conduct problems score at age 7	MCS4	4023	1.19	1.43	0	10
Conduct at Ages 11-13	SDQ conduct problems score at age 11	MCS5	4023	1.23	1.46	0	10
Conduct at Age 14+	SDQ conduct problems subscale at age 14	MCS6	4023	1.25	1.53	0	10
Impact of Child's Conduct Problems up to Age 4	SDQ impact supplement score at age 3	MCS2	4023	0.09	0.51	0	8
Impact of Child's Conduct Problems at Ages 5-6	SDQ impact supplement score at age 5	MCS3	4023	0.13	0.66	0	10
Impact of Child's Conduct Problems at Age 7+	SDQ impact supplement score at age 7	MCS4	4023	0.23	0.88	0	10
Cognitive Skills							
Cognitive Skills up to Age 4	Cognitive skills measure constructed using PCA of test scores at age 3, standardised	MCS2	4023	1.04	0.13	0.47	1.49
Cognitive Skills at Ages 5-6	Cognitive skills measure constructed using PCA of test scores at age 5, standardised	MCS3	4023	1.04	0.13	0.50	1.79
Cognitive Skills at Ages 7-10	Cognitive skills measure constructed using PCA of test scores at age 7, standardised	MCS4	4023	1.03	0.14	0.54	1.45
Cognitive Skills at Ages 11-13	Cognitive skills measure constructed using PCA of test scores at age 11, standardised	MCS5	4023	1.03	0.15	0.38	1.51
Cognitive Skills at Age 14+	Cognitive skills measure constructed using PCA of test scores at age 14, standardised	MCS6	4023	1.02	0.15	0.33	1.50
Other Child's Characteristics							
Child's Sex	Indicator if child is male	MCS1	4023	0.49	0.50	0	1
Teenage Smoking	Indicator that a child smokes at 14	MCS6	4023	0.14	0.34	0	1
Childhood Emotional Wellbeing	Indicator of SDQ emotional problems score within abnormal range	MCS3	4023	0.03	0.16	0	1
Social Conditions							
Parental Income	OECD equivalised household income after taxes and benefits, in year 2000 GBP	MCS1	4023	35,138	20,194	1,445	128,277
Parental Wealth	Parental assets, in year 2011 GBP	MCS5	4023	5,139	112,304	0	7,000,000
Parental Socio-Economic Position	Income quintile of child's household	MCS1	4023	3.26	1.32	1	5
Childhood Poverty	Indicator of child's household income below 60% median equivalised houselhold income	MCS1	4023	0.20	0.40	0	1
Parental Characteristics							
Parental Education	Indicator of parent having a university degree (NVQ 4 or above)	MCS1	4023	0.43	0.50	0	1
Parental Mental Health	Indicator of parent suffering from any diagnosed mental illness	MCS1	4023	0.08	0.27	0	1

Note: MSCj denotes data from MCS sweep j for $j = 1, 2, 3, 4, 5, 6$. Children were 9 months old during sweep 1, 3 years old during sweep 2, 5 years old during sweep 3, 7 years old during sweep 4, 11 years old during sweep 5, 14 years old during sweep 6; PCA stands for principal component analysis.

2.2 Modelling Later Life Outcomes

To model later life outcomes, we use equations parametrised using two types of evidence:

(i) target data from observational studies, which describe expected levels of and associations between variables at a point in time;

(ii) causal effect estimates from quasi-experimental studies, which attempt to draw causal inferences about the effect of one variable on another variable, either at the same time or a future point in time.

Table 2: Target Data used in Equation Modelling

Parameter	Description	Source
$\overline{dead}[age, sex, sep]$	Mortality rates by age, sex and the English Index of Multiple Deprivation 2015 (IMD) quintile group.	ONS, 2011
$\overline{depressed}[age, sex, sep]$	For ages 5-18: proportion of people with emotional problems, defined as having SDQ subscore within the abnormal range (by age, sex, and parental income quintile group); for age 18+: depression diagnosed by a doctor and present or being treated within the past 12 months (by age, sex, and English IMD quintile group).	For age 5-18: MCS; age 18+: Health Survey for England, 2014
$\overline{chd}[age, sex, sep]$	Proportion of people with CHD (by age, sex, and English IMD quintile group).	Health Survey for England, 2006
$\overline{earnings}[age, sex]$	Mean full time annual gross pay in UK (by age and sex).	Annual Survey of Hours and Earnings, ONS, 2015
$\overline{employed}[age, sex]$	Seasonally adjusted employment rate, expressed as a proportion of the economically active population (by age and sex).	Labour Force Survey, ONS
$\overline{sdq}[age, sex]$	Mean SDQ conduct problems subscale score (by age and sex).	MCS, 2000-2014
$\overline{cog}[age, sex]$	Mean cognitive measure (by age and sex).	MCS, 2000-2014
$\overline{edu}[19]$	Higher Education Initial Participation Rate in 2015/2016 (estimate of the likelihoodof a person participating in Higher Education by age 30, based on current participation rates).	Department for Education, 2016
$\overline{smokes}[14, sex]$	Proportion of 14-year-old children smoking (by sex).	MCS, 2000-2014
$\overline{smokes}[19, sex]$	Proportion of people smoking (by sex) in the age group 16-24 (proxy for age 19 in the simulation) in Great Britain.	Opinions and Lifestyle Survey, ONS, 2015
$\overline{pov}[sex]$	Proportion of households below average income (60 % median income) by sex in UK.	Family Resources Survey, Department for Work & Pensions, 2016/2017
$\overline{cd}[4, sex]$	Proportion of 4-year-old children having a conduct disorder (by sex).	The survey of the mental health of children and young people living in private households in Great Britain 2004, ONS
$\overline{prison}[age, sex]$	Average proportion of people in prison (by age and sex) in England and Wales over 31 March 2017 - 31 March 2018 (calculated using population estimates in mid-2017)	Offender Management Statistics, Ministry of Justice, 2017-2018; Population Estimates for UK, England and Wales, Scotland and Northern Ireland Mid-2017, ONS
$\overline{care}[70, sex]$	Proportion of people aged 65+ in resident care homes (by sex) in England and Wales, 2011	“Changes in the older resident care home population between 2001 and 2011” 2014, ONS

Note: ONS – Office for National Statistics, MCS – Millennium Cohort Study, U-S – Understanding Society

Our target data comes from the most up-to-date and nationally representative available surveys and administrative records in England, as do most of the causal effect estimates – with the exception of cases where more robust and relevant estimates are available from causal inference studies in other high-income countries. A full list of the target datasets is presented in table 2. Using these datasets to calibrate the trends and intercepts of our simulated variables relies on the assumption that historical cohorts are a reliable proxy for the modelled cohort.

We use four kinds of equations: (i) simple level equations based on target data only; (ii) complex level equations based on target data supplemented with causal effect estimates; (iii) simple difference equations based on target data only; (iv) complex difference equations based on target data supplemented with causal effect estimates. We illustrate each below in turn.

2.2.1 Level Equations

To model the individual probability of dying a simple approach is to use historical mortality rates from a target dataset:

$$pr.dead_{i,age} = \overline{dead}[age_i, sex_i] \tag{1}$$

where $\overline{dead}[age_i, sex_i]$ is the mean probability of dying conditional on age and sex calculated from a target dataset, the Office for National Statistics (ONS) mortality data. The bar above a variable denotes an average calculated from a target dataset.

We can also supplement equation (1) with causal effects estimates. For example, we may wish to model that coronary heart disease (CHD) increases one's probability of dying by a certain proportion, denoted $\beta_{chd}^{pr.dead}$:

$$\begin{aligned} pr.dead_{i,age} &= f(\,\overline{dead}[age_i, sex_i], \quad \overline{chd}[age_i, sex_i], \quad chd_{i,age}, \quad \beta_{chd}^{pr.dead}) = \\ = \max[0, \min[1, &\quad \overline{dead}[age_i, sex_i]\,(1 + \beta_{chd}^{pr.dead}\,(\,chd_{i,age} - \overline{chd}[age_i, sex_i]\,)\,) \quad]] \end{aligned} \tag{2}$$

where $chd_{i,age}$ is the simulated binary outcome of individual i having a CHD at a certain age; $\overline{chd}[age_i, sex_i]$ is the the mean CHD prevalence given age and sex, calculated from a different target dataset (see table 2). We subtract the target CHD prevalence from the simulated CHD outcome to avoid double counting, as the term $\overline{dead}[age_i, sex_i]$ is not independent from CHD, even if the CHD variable is not observable in the mortality target dataset.

2.2.2 Difference Equations

If a level of a variable has been modelled for the previous time periods, sometimes we proceed by modelling it's evolution as a difference equation. For example, when the level of earnings has been established at the start of 'working years' (age 19), we can model change in individual earnings during the subsequent periods as:

$$\triangle earnings_{i,age} = trend \tag{3}$$

where $\triangle earnings_{i,age} = earnings_{i,age} - earnings_{i,age-1}$ is the change in earnings from the previous year, and *trend* is a trend that governs the changes in earnings over time.

Similar to level-equations, we can supplement equation (4) with a causal effects estimate. For example, to model that developing a depression reduces earnings by a certain level, $\beta^{earnings}_{depressed}$:

$$\triangle earnings_{i,age} = trend + \beta^{earnings}_{depressed} \triangle depressed_{i,age} \tag{4}$$

where $depressed_{i,age}$ is an indicator of an individual having a depression at a given age and $\triangle depressed_{i,age} = depressed_{i,age} - depressed_{i,age-1}$

2.3 The Illustrative Policy Evaluation

We illustrate how LifeSim can be applied to evaluating a hypothetical national parent training programme. We make a simple comparison between delivering publicly funded parent training to all eligible parents versus none. We take short-term effect data from a recent systematic review of evidence about the effects of one particular parent training on child conduct problem – the "Incredible Years" programme (Gardner et al., 2017).

We assume that the hypothetical policy intervention:

(i) is delivered to parents of all five-year old children screened as being at risk of developing a conduct disorder, based on a parent-reported SDQ conduct problems score at age 5 within the abnormal range (4 or above);

(ii) causes an average 0.46 standard deviation decrease in the SDQ conduct problems score of a child recipient, with heterogeneous effects conditional on child and parental characteristics (larger effects for the children of parents with mental health problems and for children with a higher baseline conduct problems score, and correspondingly smaller effects for other children (Gardner et al., 2017)).

We also conduct sensitivity analysis using alternative assumptions about effectiveness, including a random error reflecting individual heterogeneity and a conservative assumption of effect fade-out over time (Feinstein et al., 2017; van Aar et al., 2017) (see Supplementary Appendix C).

3 Results

Table 3 summarises the simulated benefits of the illustrative parental training programme for the child recipients.

Table 3: Policy Benefits for the Recipient Children

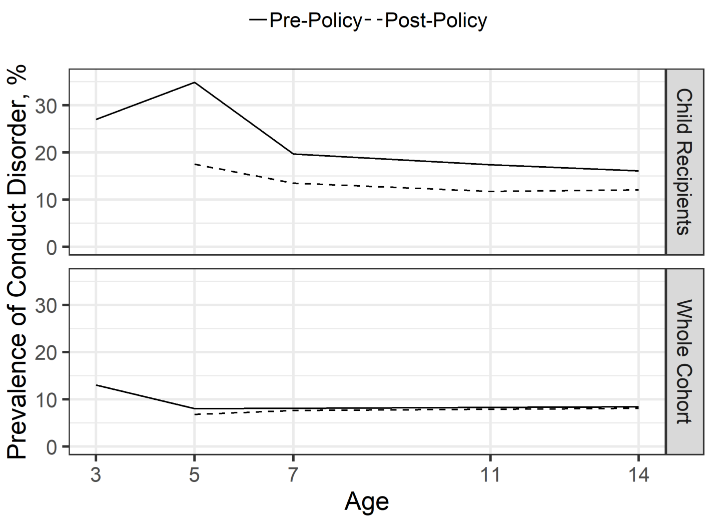
Outcome	Pre-Policy	Post-Policy	Absolute Effect	Relative Effect, %
Child Outcomes				
Conduct Disorder, Age 5, %	35.64	17.68	-17.96	-50.39
Conduct Disorder at Age 18, %	16.05	11.95	-4.11	-25.59
SDQ Conduct Problems Score at Age 5	4.64	3.99	-0.65	-12.21
SDQ Conduct Problems Score at Age 18	2.87	2.28	-0.59	-8.26
Cognitive Skills at Age 5 (standardised)	1.00	1.00	0.00	0.00
Cognitive Skills at Age 18 (standardised)	0.99	0.99	0.00	0.00
Adult Outcomes				
University Graduates, %	42.06	42.76	0.70	1.66
Working Years in Unemployment, %	8.85	7.61	-1.24	-14.00
Life Years in Poverty, %	36.37	35.77	-0.61	-1.67
Working Years in Prison, %	2.82	2.10	-0.71	-25.30
Retirement Years in Residential Care, %	3.59	2.93	-0.66	-18.32
Adult Years as a Smoker, %	33.46	31.82	-1.64	-4.91
Adult Years with CHD, %	5.96	5.98	0.02	0.25
Life Years with Mental Illness, %	9.97	8.36	-1.61	-16.11
Years of Life	79.17	79.27	0.09	0.12
Premature Mortality Rate (before age 75), %	28.45	28.20	-0.25	-0.88
Final Wellbeing Outcomes				
Annual Consumption (lifetime average), £	16,981	17,004	22.40	0.13
Years of Healthy Life	75.74	76.14	0.40	0.53
Years of Healthy Life (discounted)	44.52	44.68	0.16	0.37
Years of Good Life	71.30	71.72	0.41	0.58
Years of Good Life (discounted)	41.32	41.49	0.17	0.42

Note: The effects are calculated on average per child recipient (7,166 child recipients in total). CHD – coronary heart disease; SDQ conduct problems score ranges 0-10 with a higher value representing more conduct problems; cognitive skills measure is a common factor extracted from the cognitive skills measures disseminated by MCS, with a higher value representing better skills, standardised with a mean of 1.00 and standard deviation of 0.15. Years of healthy and good life are discounted at 1.5% annually.

Table 3 shows that the assumed decrease in the SDQ conduct problems score at age 5 corresponds to an absolute decrease in the score of around 0.65 points on average, which is a 12.2% relative decrease. This translates into preventing around 17.0% of the children from developing conduct disorder at the age 5.

However, the effect diminishes by age 18, with only 4.1% of conduct disorder cases being prevented by age 18. This occurs because most children with high parent-reported conduct problem scores in the early years progress to scores within the normal range in later childhood (what we might call "progression to the mean"), while some deteriorate to even worse scores. This gradual polarisation of conduct problem scores causes a natural policy effect fade-out – the initial effect is large, as conduct scores fall more rapidly than they would otherwise have done, but the effect then gradually reduces as many children go through the natural process of improvement in parent-reported conduct score (see Figure 3).

Figure 3: Prevalence of Conduct Disorder over Time, Pre and Post Policy-Policy Intervention



Note: The above shows the prevalence conduct disorder for children at different ages before and after the policy intervention. The top panel depicts the prevalence among the 7,166 child recipients, the bottom panel – prevalence in the whole cohort of 100,000 individuals.

The policy also has effects on the proportion of university graduates (0.7% point increase), proportion of adult years unemployed (1.2% point reduction), proportion of years in poverty (0.6% point reduction) proportion of working years in prison (0.7% point reduction), proportion of retirement years in residential care (0.7% point reduction), proportion of adult years as a smoker (1.6% point reduction), proportion of adult years with mental illness (1.6% point reduction).[2]

Finally, the intervention increases the average annual consumption of child recipients by around £22, and lifetime health by 0.40 years of healthy life. These two outcomes, when combined in a wellbeing measure that adjusts the quality of life for health quality and consumption together (Cookson et al., 2016), translate into a 0.41 point increase in the total lifetime years of good life.

Table 4 reports the estimated public budget savings generated by the intervention (discounted at 1.5% annual rate, see Paulden and Claxton (2012)). There are substantial initial savings due to reduced costs to social, educational and health services for children with conduct disorders, with further savings in adulthood due to reduce costs to the criminal justice system, additional tax revenues and lower benefit payments. The cost of the "Incredible Years" falls within the range £1,612-2,418 per recipient, depending on the training group size (Edwards et al., 2016). This implies that the initial savings would cover the costs of the programme within a five to ten year period, with further public cost savings in adulthood. There are also savings in terms of increased tax revenues and decreased benefit payments. The total government budget savings per recipient sum up to £1,177 within the first five years and £16,353 over lifetime.

Table 4: Cost Savings as a Result of the Policy Intervention

Public Cost Savings and Revenues (per recipient), £	5 years	10 Years	15 Years	20 Years	Lifetime
Conduct Disorder	1,051	1,412	1,511	1,511	1,511
Healthcare: CHD	0	0	0	0	-3
Healthcare: Mental Illness	125	192	242	290	2,444
Prison	0	0	197	1,110	6,624
Residential Care	0	0	0	0	681
Benefit Payments	0	0	71	434	3,721
Tax Revenues	0	0	0	9	1,374
Total Savings	1,177	1,604	2,020	3,355	16,353

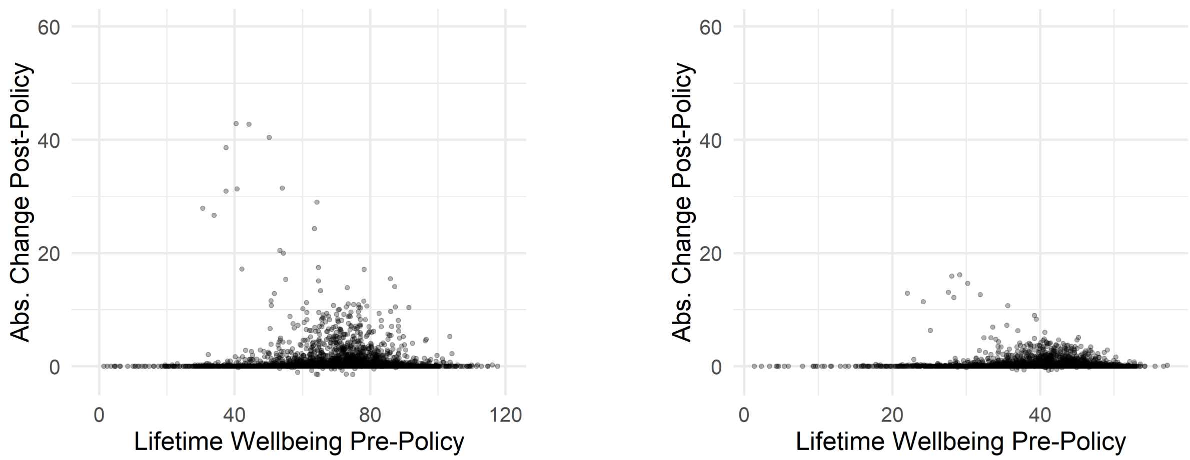
Note: Savings estimated per young child at risk of conduct disorder at age 5, and discounted at 1.5 % annual rate. The unit conduct disorder costs are estimated by Bonin et al. (2011) for the annual cost of a child with a conduct disorder, and include costs to NHS, Social Services, Education and Voluntary sector. See details on other cost sources in the Supplementary Appendix D. The estimated cost of the "Incredible Years": £1,612-2,418 per child recipient (Edwards et al., 2016).

[2] There is a small and counter-intuitive increase the proportion of life years spent with CHD. This occurs in people who shift to a situation of job market instability, rather than long-term unemployment, which we assume has a negative effect on mental health, smoking and, thus, CHD. This is because we find robust quasi-experimental evidence that becoming unemployed has a negative impact on mental health which is not fully offset by becoming employed again, but no robust data on the differential mental health effects of short-term versus long-term unemployment. Rather than applying an arbitrary fix, however, we prefer to present full details of our findings, warts and all.

Furthermore, Figure 4 shows that even though the model predicts small average lifetime benefits of the parent training programme, some individuals experience substantial benefits – for example, 143 children (2.00% of the 7,166 recipients) gain 5 or more good years of life, and 39 children (0.54% of the 7,166 recipients) gain even 10 or more years of good life.

Figure 4: Policy Gains in Terms of Lifetime Wellbeing: by the Pre-Policy Wellbeing Level

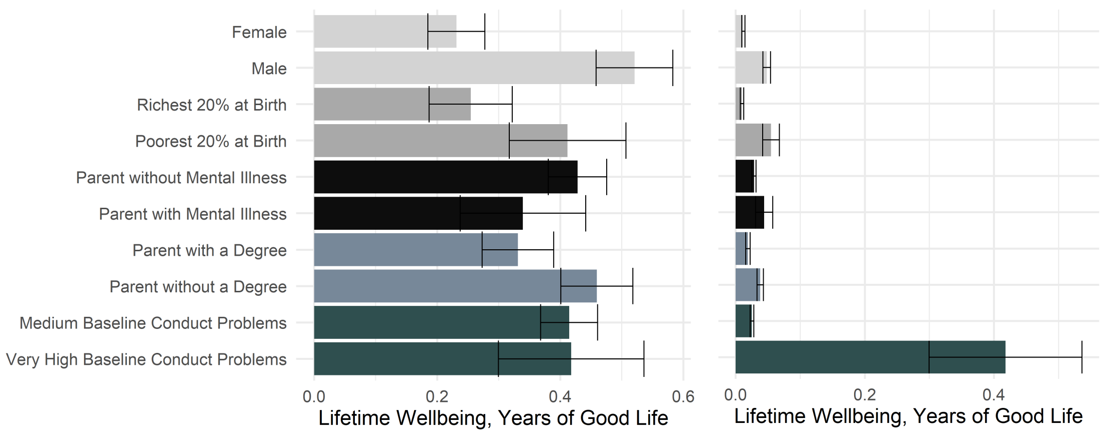
(a) *no discount rate* (b) *discounted at 1.5 % annually*



Note: The above represents the policy gain in terms of good years of life for the 7,166 child recipients, given the good years of life pre-policy.

Figure 5: Distributional Effects: Average Policy Gains in Terms of Wellbeing

(a) *the groups of recipients only* (b) *the entire cohort*



Note: The above represents the average gain in terms of lifetime wellbeing for the subgroups; (a) represents the impact among the group of 7,166 recipients; (b) represents the impact in the entire cohort. The black line represents 95 % confidence interval. High baseline conduct problems are defined by SDQ conduct problems score (at age 5) ≥ 6; medium baseline conduct problems are defined by SDQ conduct problems score (at age 5) < 6 and ≥ 4.

Figure 5 demonstrates how the framework can be useful in equity analysis. The left hand panel looks at inequality within the population of recipients, and the right hand panel looks at inequality within the general population. Each bar represents the average gain in wellbeing for a subgroup characterised by one of the following: sex, parental income quintile group, parental mental health, parental education, and the baseline conduct problem score at age 5. The left hand panel shows that the intervention has a larger impact on improving wellbeing in more disadvantaged recipient children according to few indicators of social disadvantage (the poorest 20 % at birth, children whose parents are without a university degree).[3] In addition, boys gain more from the intervention, a finding consistent with previous literature (Gardner et al., 2017). The right-hand panel examines equity impacts in the general population, allowing for differences in the number of recipients and non-recipients within each social group. At a general population level, the policy also appears equity-improving, with particular benefits to children from the poorest 20% of households at birth, a parent with mental illness, a parent without a degree, and with very high baseline conduct problems.

Table 5: Whole-cohort Lifetime Inequality Impacts: Impact by Childhood Circumstance

Childhood circumstance	Annual Consumption, £			Lifetime Health, Years of Healthy Life			Lifetime Wellbeing, Years of Good Life		
	Abs. Gain	95% CI		Abs. Gain	95% CI		Abs. Gain	95% CI	
Female	0.08	0.03	0.14	0.01	0.01	0.01	0.01	0.01	0.01
Male	3.42	2.92	3.91	0.05	0.04	0.05	0.05	0.04	0.05
Richest 20% at Birth	0.46	0.22	0.71	0.01	0.01	0.01	0.01	0.01	0.01
Poorest 20% at Birth	3.01	1.77	4.26	0.05	0.04	0.07	0.05	0.04	0.07
Parent without Mental Illness	1.69	1.43	1.95	0.03	0.02	0.03	0.03	0.03	0.03
Parent with Mental Illness	2.04	1.35	2.74	0.04	0.03	0.06	0.04	0.03	0.06
Parent with a Degree	1.10	0.83	1.36	0.02	0.01	0.02	0.02	0.02	0.02
Parent without a Degree	2.19	1.81	2.57	0.04	0.03	0.04	0.04	0.03	0.04
Medium Baseline Conduct Problems	1.41	1.19	1.64	0.02	0.02	0.03	0.03	0.02	0.03
Very High Baseline Conduct Problems	27.97	19.64	36.30	0.40	0.28	0.52	0.42	0.30	0.54
Best Off Extreme Group	1.19	0.74	1.64	0.02	0.02	0.03	0.02	0.02	0.03
Worst Off Extreme Group	39.95	13.73	66.17	0.35	0.18	0.53	0.38	0.19	0.57
Best Off 20%	0.51			0.01			0.01		
Worst Off 20%	2.76			0.06			0.06		

Note: The above represents the average gains per cohort member for the subgroups of the simulated cohort of 100,000 individuals. The "Best Off" and "Worst Off" subgroups are calculated based on all of the childhood circumstances above, except for sex.

Table 5 provides more detail on the general population inequality impacts, in terms of annual

[3] Even though the policy has a higher short-term impact on the conduct problems among those children who have a parent with worse mental health, the long term wellbeing benefits are higher for the children whose parents do not have a mental illness.

consumption, total healthy and good life years. It also provides two summary measures of impact on inequality in lifetime wellbeing, based on differences in wellbeing between best off and worst off groups. Our "best off extreme group" focuses on individuals in the top category of all four main markers of social disadvantage in early life (top 20% parental income, high parental education, no parental mental illness, high baseline conduct disorder problems). Our "best off 20%", by contrast, focuses on the best off 20% of individuals in terms of predicted lifetime wellbeing based on all four main markers of social disadvantage in early life.

4 Discussion

4.1 Summary of Main Findings

We find that an illustrative parent-training programme can substantially reduce cases of conduct disorder from age 5 to 7, though the effect partially fades out since many conduct problems would have resolved by age 11 in any case. Despite this fade out, we estimate that public cost savings are sufficient to cover the cost of the programme within the first 5 years and that further public savings accrue into adulthood.

While later life benefits are small for most intervention recipients, a subset of recipients gain substantial long-term benefits in terms of both length and quality of life – including better material living standards as well as reduced risks of mental and physical illness, poverty, unemployment and imprisonment. Our subgroup analysis suggests that these substantial beneficiaries are disproportionately children from socially disadvantaged backgrounds, and that the programme contributes to reducing inequality of opportunity for lifetime wellbeing.

The prevalence of conduct disorder falls with age among the at risk children targeted for intervention, whilst rising in the cohort as a whole. This suggests that some children who could have benefited from the intervention were not picked up by the screening process as their conduct disorder was not apparent at the screening age but only developed later in their childhood (see the policy implications in section 4.3).

4.2 Strengths and Limitations

The key strength of our approach is that it captures the clustering and compounding of disadvantage over the lifecourse and how this generates substantial individual-level heterogeneity in policy outcomes. Poor conduct in early years results in poor educational and employment outcomes, sometimes resulting in spells in prison and often leading to poor health, which then manifests itself in costs to public services and the social protection system.

Another advantage of our approach is that it can integrate the various outcomes that occur over the lifecourse into an overarching concept of lifetime wellbeing.

These two features allow us to provide more detailed and accurate information about the long-term benefits of social policies where more traditional modelling approaches would: (i) assume benefits in one stage in life do not impact later life outcomes; (ii) assume these independent impacts at different life stages accrue to different individuals; (iii) have no way of aggregating the different benefits across the different types of outcomes.

Our prototype model has various limitations.

We have devoted considerable time and effort to parameterising each of our lifecourse equations by reviewing literature and consulting experts to identify the best available evidence, but larger teams of researchers would be able to improve each equation by adopting more systematic approaches to reviewing evidence and eliciting expert beliefs about biases in applying causal effects estimated using historical data to predicting future trends. For example, our validity checks found implausibly low predicted rates of mental illness and smoking in adulthood.

As with any modelling study we have had to select which aspects of reality to capture in the model and which aspects to omit. For example, our prototype model leaves out many social and biological variables which in a more comprehensive model would act as independent, mediating and moderating factors within our complex network of causal pathways, such as:

- social environment (e.g. relationship quality, childhood attachment to parents, friends and school, loneliness, parenting quality, social support, and social networks)
- race/ethnicity
- geographical environment (e.g. neighbourhood quality, urban-rural differences, regional differences)

- adverse childhood experiences (such as neglect, violence and family breakdown),
- other adverse adulthood experiences (e.g. debt, homelessness)
- other unhealthy behaviours (e.g. poor diet, physical inactivity, drug addiction)
- health capital (including genetic factors such as gestational age, birth weight and birth defects as well as biomedical factors such as adiposity and blood pressure)
- biological indicators of psycho-social stress (e.g. cortisol levels, telomere length), including epi-genetic factors involving heritable changes in gene function.

Our summary measure of wellbeing focuses only on health outcomes, consumption and mortality, without incorporating any independent wellbeing impact of adverse experiences such as unemployment and imprisonment, and requires explicit value judgements that have not (yet) been extensively tested in value elicitation studies. The model also make a number of simplifying assumptions. For example, health outcomes are modelled using just three binary variables – mental illness (depression), physical illness (CHD) and mortality – educational outcomes focus only on gaining a university degree; employment outcomes focus only on unemployment not precarious employment; and our modelling of the tax and benefit system and retirement savings is extremely stylised.

Finally, our prototype model has more fundamental limitations. We only simulate a single birth cohort rather than the all-age population, we model public costs in a simple way that ignores sector-specific budget constraints and opportunity costs, we do not model the dynamics of family formation and dissolution or spillover effects on other family members, and we do not model behavioural responses to policy change or general equilibrium effects. All of these limitations can potentially be addressed within our general framework by improving our prototype model.

4.3 Policy Implications

Our findings are consistent with studies finding substantial public cost savings from parenting programmes, such as Bonin et al. (2011), who estimate that the "Incredible Years" programme generates cost savings within the first 5-8 years. However, they also highlight the importance of accurate targeting to families who can benefit the most. There may be a trade-off between delivering programmes at an earlier age when the potential benefits are greater (List et al., 2018), versus delivering them at a later age when problems can be more accurately diagnosed and the programme more effectively targeted.

4.4 Research Implications

Despite the many limitations of our prototype microsimulation model, the general framework within which it is embedded is extremely flexible and opens up an exciting research agenda for lifecourse economic evaluation. Policy-makers are often accused of “short-termism”, and the lifecourse perspective often receives short shrift in public debates. Lifecourse economic evaluation can potentially help keep the lifecourse perspective in view, by routinely providing policy makers with detailed and credible information about lifecourse policy consequences. We hope that this prototype study will encourage others to develop better methods of lifecourse economic evaluation, which address the many limitations of our prototype model and provide policy makers with useful insights about the lifecourse consequences of alternative policy options across all sectors of public policy.

Acknowledgements

We would first like to thank the members of our advisory group: Annalisa Belloni, Sarah Cattan, Leon Feinstein, Paul Frijters, Peter Goldblatt, Heather Joshi, Catherine Law, Lara McClure, Mark Petticrew and Christine Power. For useful conversations and comments we also are grateful to Shehzad Ali, Karen Bloor, Laura Bojke, Eva Maria Bonin, Jonathan Bradshaw, Tracey Bywater, Simon Capewell, Maria Guzman Castillo, Bette Chambers, Brendan Collins, Tim Doran, Susan Griffin, Nils Gutacker, Andrew Jones, Noemi Kreif, Christodoulos Kypridemos, Richard Mattock, Cheti Nicoletti, Martin O’Flaherty, Kate Pickett, Gerry Richardson, Jemimah Ride, Matthew Robson, Tracey Sach, Tushar Srivastava, David Taylor-Robinson, Valentina Tonei, Aki Tsuchiya, Simon Walker and Margaret Whitehead.

References

Bonin, E.-M., Stevens, M., Beecham, J., Byford, S., and Parsonage, M. (2011). Costs and longer-term savings of parenting programmes for the prevention of persistent conduct disorder: a modelling study. *BMC public health*, 11(1):803.

Cookson, R., Cotton-Barrett, O., Adler, M. D., Asaria, M., and Ord, T. (2016). Years of good life based on income and health: Re-engineering cost-benefit analysis to examine policy impact on wellbeing and distributive justice.

Doyle, O., Harmon, C. P., Heckman, J. J., and Tremblay, R. E. (2009). Investing in early human development: timing and economic efficiency. *Economics & Human Biology*, 7(1):1–6.

Edwards, R. T., Jones, C., Berry, V., Charles, J., Linck, P., Bywater, T., and Hutchings, J. (2016). Incredible years parenting programme: cost-effectiveness and implementation. *Journal of Children's Services*, 11(1):54–72.

Feinstein, L., Chowdry, H., and Asmussen, K. (2017). On estimating the fiscal benefits of early intervention. *National Institute Economic Review*, 240(1):R15–R29.

Frijters, P., Bellet, C., and Krekel, C. (2017). Micro-macro simulations for wellbeing, version 1.0. *Working Paper*.

Garces, E., Thomas, D., and Currie, J. (2002). Longer-term effects of head start. *American economic review*, 92(4):999–1012.

García, J. L., Heckman, J. J., Leaf, D. E., and Prados, M. J. (2017). Quantifying the life-cycle benefits of a prototypical early childhood program. Technical report, National Bureau of Economic Research.

Gardner, F., Burton, J., and Klimes, I. (2006). Randomised controlled trial of a parenting intervention in the voluntary sector for reducing child conduct problems: outcomes and mechanisms of change. *Journal of child psychology and psychiatry*, 47(11):1123–1132.

Gardner, F., Leijten, P., Mann, J., Landau, S., Harris, V., Beecham, J., Bonin, E.-M., Hutchings, J., and Scott, S. (2017). Could scale-up of parenting programmes improve child disruptive behaviour and reduce social inequalities? using individual participant data meta-analysis to establish for whom programmes are effective and cost-effective.

Goodman, R., Ford, T., Simmons, H., Gatward, R., and Meltzer, H. (2003). Using the strengths and difficulties questionnaire (sdq) to screen for child psychiatric disorders in a community sample. *International Review of Psychiatry*, 15(1-2):166–172.

Goodman, R., Renfrew, D., and Mullick, M. (2000). Predicting type of psychiatric disorder from strengths and difficulties questionnaire (sdq) scores in child mental health clinics in london and dhaka. *European child & adolescent psychiatry*, 9(2):129–134.

Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science*, 312(5782):1900–1902.

Heckman, J. J., Moon, S. H., Pinto, R., Savelyev, P. A., and Yavitz, A. (2010). The rate of return to the highscope perry preschool program. *Journal of public Economics*, 94(1-2):114–128.

Hills, J. (2017). *Good Times, Bad Times (revised Edition): The Welfare Myth of Them and Us*. Policy Press.

Hutchings, J., Bywater, T., Daley, D., Gardner, F., Whitaker, C., Jones, K., Eames, C., and Edwards, R. T. (2007). Parenting intervention in sure start services for children at risk of developing conduct disorder: pragmatic randomised controlled trial. *BMJ: British Medical Journal*, 334(7595):678.

Joshi, H., Nasim, B., and Goodman, A. (2016). The measurement of social and emotional skills and their association with academic attainment in british cohort studies. In *Non-cognitive Skills and Factors in Educational Attainment*, pages 239–264. Springer.

Layard, R., Clark, A. E., Cornaglia, F., Powdthavee, N., and Vernoit, J. (2014). What predicts a successful life? a life course model of wellbeing. *The Economic Journal*, 124(580):F720–F738.

Lee, S., Aos, S., Drake, E., Pennucci, A., Miller, M., and Anderson, L. (2012). Return on investment: Evidence-based options to improve statewide outcomes, april 2012 (document no. 12-04-1201) washington state institute for public policy. *Olympia, WA.*

List, J. A., Samek, A., and Suskind, D. L. (2018). Combining behavioral economics and field experiments to reimagine early childhood education. *Behavioural Public Policy*, pages 1–21.

Paulden, M. and Claxton, K. (2012). Budget allocation and the revealed social rate of time preference for health. *Health Economics*, 5(21):612–618.

Paull, G. and Xu, X. (2017). Study of early education and development (SEED): The potential value for money of early education. *Frontier Economics Research Report for the Department of Education.*

Schweinhart, L. J., Montie, J., Xiang, Z., Barnett, W. S., Belfield, C. R., and Nores, M. (2005). Lifetime effects: The high/scope perry preschool study through age 40 (monographs of the high/scope educational research foundation, 14). *Ypsilanti, MI: High Scope Educational Research Foundation.*

Scott, S., Briskman, J., and O'Connor, T. G. (2014). Early prevention of antisocial personality: long-term follow-up of two randomized controlled trials comparing indicated and selective approaches. *American Journal of Psychiatry*, 171(6):649–657.

van Aar, J., Leijten, P., de Castro, B. O., and Overbeek, G. (2017). Sustained, fade-out or sleeper effects? a systematic review and meta-analysis of parenting interventions for disruptive child behavior. *Clinical psychology review*, 51:153–163.

Zucchelli, E., Jones, A. M., and Rice, N. (2012). The evaluation of health policies through dynamic microsimulation methods. *International Journal of Microsimulation*, 5(1):2–20.