Department for Education

Schools policy appraisal handbook

Research report

June 2021

Louis Hodge, Allan Little and Matthew Weldon: Department for Education

Government Social Research

Contents

List of Figures

List of Tables

Introduction

This Handbook is a companion to our report on *GCSE Attainment and Lifetime Earnings* (Hodge, Little and Weldon (2021)). Wages are commonly used to value the economic benefits of investments in education. We are motivated by the need to use these estimates appropriately, in the appraisal and evaluation of school-based policy interventions.

There is no perfect or comprehensive way to assess the social value of education policy. In cost-benefit analysis (hereafter, CBA), we subscribe to the statistician George Box's oft-quoted aphorism that *"all models are wrong, but some models are useful"* (Box, Luceño and del Carmen Paniagua-Quinones (2009)). Box (1976) outlined a process to improve statistical models, by iterating between the theory and the practice, with selective concern for key model limitations. Similarly, Dreze and Stern (1987) called for researchers involved in developing CBA models to bring "the difficulties of the practitioner to greater prominence in the theory".

The Handbook engages in a short series of theory-practice iterations, with a particular focus on how earnings enter cost-benefit frameworks for schools policy. We wish to engage in this process in an open, transparent and deliberative manner. The Handbook aims to be enabling, rather than prescriptive - it does not give unconditional endorsement to a singular approach.

Sections of the Handbook can be read independently, as a series of vignettes. We discuss the theory, estimation and suggested practice, themed around:

1. **Social welfare**: A basic CBA might sum individual earnings benefits on one side of the ledger and then subtract the costs of investment. In this form, the CBA is narrow and makes overly strong welfare assumptions. We consider the use of "life course" models of subjective wellbeing, which can be adapted to describe the outcomes of most policy interventions for children and young people. This would help to set any subsequent CBA metrics into context.
2. **Distributional weights**: Income improves wellbeing at a decreasing rate (Layard, Nickell and Mayraz (2008)). A higher weighting factor could be placed on additional earnings, accruing to people on lower incomes. Even with uncertainty, weighting has the potential to transform education policy appraisals. Weighting can offset some structural biases, given that data on earnings can imply a lower marginal value of education for groups that have historically earned less, e.g. women, pupils from disadvantaged backgrounds and pupils in low-income areas.
3. **Productivity**: We recommend applying a productivity 'uplift' to earnings, to represent employers' willingness to pay more highly-skilled workers. We discuss the merits of different levels of adjustment.

4. **Growth**: Estimates of lifetime earnings are typically obtained from microeconomic models. Macroeconomic models provide an alternative, capturing the wider benefits of education across the whole economy. We discuss the pros and cons of these two approaches in CBA.
5. **Signalling**: Test scores can reflect skills acquired during school but also act as a signal of pupils' existing abilities. We note the consensus that investments in education predominantly add to human capital. On balance, earnings are more likely to under-value than over-value the productivity benefits of educational investments. However, signalling is context-specific and its implications ought to be considered, in each appraisal.
6. **Wellbeing**: Wellbeing research can be used at all stages in the policy-making cycle, from formulating options through to evaluation. We discuss methods by which to appraise the impacts of education policy on wellbeing, offering a different perspective to models that link investment in education to wages, productivity or growth.

The Handbook is predominantly for CBA practitioners and so we assume familiarity with the principles and processes of economic appraisal, outlined in the HM Treasury "Green Book" (2020). The Green Book recommends a Five Case model covering: Strategic; Economic; Financial; Commercial and Management cases. It notes the importance of a clear policy objectives and rationale for intervention, from the outset. In the Economic Case, it is necessary to undertake a long-list appraisal, before quantifying costs and benefits in detail at the short-list stage. Further, there are supplementary Green Book guides designed to support the assessment of policies from a range of perspectives, a key example being the Enabling Natural Capital Approach (ENCA) to consider policy impacts on the natural environment. Our discussion effectively focuses on the short-list appraisal stage, taking the need for a more holistic approach to public value as a given.

Welfare

Non-technical summary

Practitioners should avoid simply adding up predicted changes in earnings, on the benefits side of the ledger, and then subtracting policy costs. In this form, CBA strays too far from the overriding objectives of investment in education, to improve people's lives.

A suitable starting point is to describe a *theory of change* for each policy, initially, without regard for whether changes in the policy inputs, outputs or outcomes can be quantified. We consider the use of "life course" models of subjective wellbeing, a flexible framework that could be adapted to describe the outcomes of most policy interventions for children and young people. We discuss that this framework is broadly compatible with insights from the the vast human capital branch of the economics literature, which underpins the CBA approach in education. Developing a logic map based on subjective wellbeing would place any CBA metrics in context, giving a clearer understanding of what they can and cannot capture. From this perspective, we can also consider ways in which to improve and adjust the CBA metrics themselves, the subject of our subsequent vignettes.

Social welfare function

Economic appraisals are founded on the principles of welfare economics. The relationship between income and social welfare can be represented as:

$$W = u_1(y_1) + u_2(y_2) + \cdots + u_n(y_n) \tag{1}$$

where W = social welfare; u_i = the utility of individual i; and y_i = the income of individual i.[1]

The metric for appraising education policy can be taken to be the effect on aggregate welfare, ΔW. This can be expressed in terms of changes in income, and approximated by:

$$\Delta W \approx \Delta y_1 \cdot u_1'(y_1) + \Delta y_2 \cdot u_2'(y_2) + \cdots + \Delta y_n \cdot u_n'(\Delta y_n) \tag{2}$$

If education policies have a net positive impact on welfare then they are worthwhile undertaking.[2] Budgets are inevitably constrained and so preference is given to policies

[1] This generic format is known as the Bergson-Samuelson social welfare function.
[2] In Utilitarian terms.

that generate the largest positive change in welfare.

In line with this theoretical model, a basic CBA for a school-based policy intervention might estimate the net change in welfare by: (i) summing the earnings benefits accruing to pupils over their lifetime; and then (ii) subtracting the costs of the policy to the taxpayer. In this format, the CBA makes overly-strong assumptions, including that:

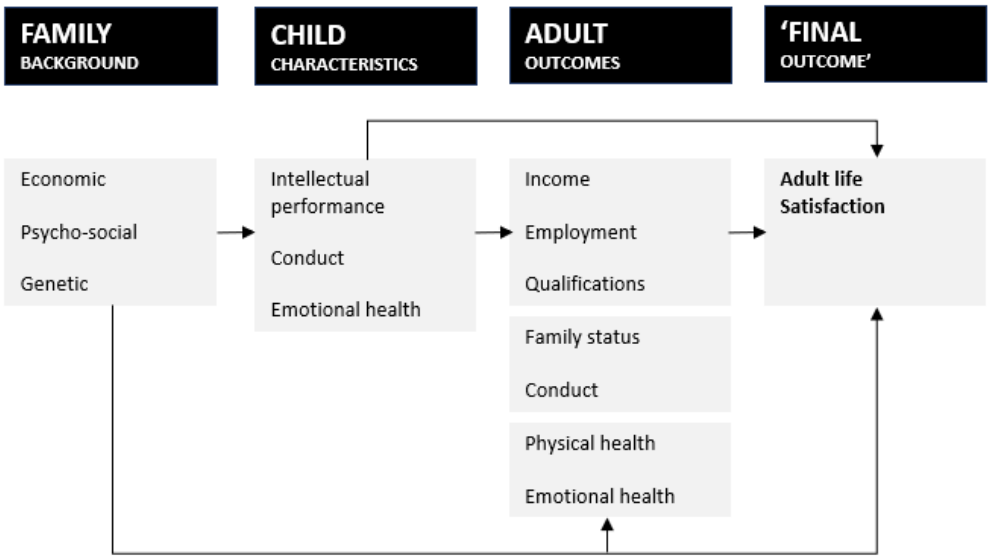
- Welfare is only determined by changes in income
- One pound of additional income has the same value for all individuals
- There is no utility interdependence between each individual, resulting in wider impacts for society

The rest of the Handbook considers ways in which appraisals might relax these assumptions. To begin that process requires a broader framework, describing the relationships between attainment, income and welfare alongside a wider set of social impacts linked to education.

Life-course Models

We might consider the impact of educational attainment from a life-course perspective, and with a broad conception of how this affects outcomes throughout people's lives. An example is given in figure 1 from O'Donnell et al. (2014), in which adult life satisfaction is treated as the outcome of interest.

Figure 1: Model of adult life satisfaction



Similar frameworks usefully breakdown phases of childhood and adolescence (see Reynolds et al. (2017); Paull and Xu (2017)). In our vignette on Wellbeing we discuss this

empirical evidence underpinning the life-course model in more detail. For example, Layard et al. (2018) estimate that the strongest childhood predictor of a satisfying adult life is emotional health in childhood, followed by behavioural outcomes, and only then by cognitive attainment. One would ideally assess the impact of each policy on all three of these outcomes, however, we are often constrained in CBA to measurable impacts on attainment, for which robust data is more readily available.

Education policy cannot easily influence all determinants of life satisfaction, notably family background and genetics: between a third and a half of the variation in wellbeing within a population has been attributed to genetic makeup (Diener and Lucas (1999); Frijters, Johnston and Shields (2013)). We do not explore this literature further, but there are two insights we ought to draw for CBA: first, that the limits of schools' influence on outcomes is constrained and so this will be reflected in the realistic objectives of the policy; and second, that policy impact evaluations ideally need to control for these background characteristics.

Developing a bespoke logic map for each intervention - perhaps taking the the life course model as an initial starting point - would usefully set the parameters for any subsequent CBA. If the CBA does focusing largely on test scores and earnings then it is liable to miss other important outcomes. Again, this is not to suggest that the CBA is not instructive, but it may speak more directly to labour market benefits and require other forms of evidence to be captured elsewhere in the appraisal.

Human capital frameworks are also commonly used to develop evidence on the economic benefits of education. This approach views the formation of skills as a production process, with inputs and outputs (see Cunha and Heckman (2008); Cunha, Heckman and Schennach (2010))

The OECD defines human capital as:

> "the knowledge, skills, competencies and attributes embodied in individuals that facilitate the creation of personal, social and economic wellbeing" (Keeley (2007))

Adopting this definition, human capital models are compatible with the life-course models of wellbeing, above. It would be possible to blend the insights from both branches of the literature, when developing a theory of change for a given policy.

Issues

We have briefly discussed life-course and human capital models, two leading candidates from which to develop bespoke theories of change for education policy. Both are

grounded in large, robust evidence and have the flexibility to describe the impacts of school-based interventions.

Even these logic models are partial and selective and so they cannot claim to be entirely free from subjective (and potentially biased) choices by the researcher. For example, we need to consider whether human capital and/or subjective wellbeing frameworks allow for interdependence between individuals. Do they allow one person's educational attainment to affect another person's outcomes?

Both branches of the literature have much to say about interdependence, and second order effects. Human capital models emphasise that people acquire knowledge from each other and skills can drive technological growth across the whole economy. The wellbeing literature highlights that an individual's relative socio-economic position matters (Di Tella and MacCulloch (2010); Clark and Oswald (2002); Easterlin (1974)). It is possible that increases in income for one person could reduce the sense of wellbeing for another, if their income does not also rise (Layard, Nickell and Mayraz (2008)). Most education policy is targeted in some way and so these secondary effects (positive and negative) are likely.

If the literature is sufficiently rich, in this regard, then the theories of change for a given policy still need to be parsimonious, inevitably overlooking some of the more complex interactions, between policies and between people. Nevertheless, a life course model can convey the most important channels, through which an education intervention affects people's lives. It sets the parameters for what can and cannot be measured in any subsequent CBA. In our case, this would put any first order estimates of the attainment-earnings route to policy impact into context.

Distributional effects

Non-technical summary

Income generally improves wellbeing at a decreasing rate (Layard et al. (2018)). That is, the value of an extra pound is worth more to people on lower incomes than to those on higher incomes. HM Treasury (2020) proposes the use of 'welfare weights' to better represent this relationship in cost-benefit analysis (CBA).

Whilst weighting is an established practice in the appraisal of employment programmes, it is rarely used in education policy appraisal. For adults, information is readily available on income to estimate weighting factors. Whereas for interventions in schools, it is necessary to predict where pupils might fall in the income distribution, later in their life.

Weighting could prove transformational in the appraisal of education policy. Given that both intergenerational social and income mobility is relatively low in the UK[3], it is likely to make a considerable difference to the outcome of any CBA, leaning toward targeted policies for disadvantaged pupils and deprived areas. This is particularly important because we find that the earnings returns to attainment are, on average, smaller for more disadvantaged groups (Hodge, Little and Weldon (2021)).

In this section we first briefly describe the technical underpinnings. Then give a short illustrative example, which shows that weighting would make a considerable difference to the implications of the CBA, leaning toward targeted policies for disadvantaged pupils and deprived areas. We then make tentative suggestions on the weighting values that could be used by practitioners.

Any use of distributional weights should be proportionate and we do not advocate their use in all cases. Whilst weighting can be informative in many cases, any analysis should be mindful of the uncertainty weights introduce. The Green Book (2020) strongly recommends that any analysis using welfare weights should be presented side-by-side with the analysis without welfare weights.

Welfare weights

[3]Gregg et al. (2017) find high income persistence across generations. Income elasticities (0.55) and rank-order correlations (0.354) are found to be high in Britain, both in absolute terms and relative to some other countries, for example Sweden.

The elasticity of marginal utility

The first step to estimating weighting factors is to calculate the *elasticity of marginal utility*, that is how much the utility value of an extra pound varies with income.

If we assume a standard isoelastic utility function:

$$u(y) = \begin{cases} \frac{y^{1-\rho}-1}{1-\rho} & \rho \neq 1 \\ ln(y) & \rho = 1 \end{cases} \tag{3}$$

Then the marginal utility of income is:

$$u'(y) = y^{-\rho} \tag{4}$$

And so the ratio between two marginal utilities is:

$$\frac{u'(y_1)}{u'(y_2)} = \left(\frac{y_2}{y_1}\right)^{\rho} \tag{5}$$

Therefore if $\rho = 1$ then marginal utility would be inversely proportional to earnings (y). For example, an individual earning £25,000 would gain twice as much utility from an extra pound than someone earning £50,000.

Most evidence suggests that the elasticity of marginal utility is greater than one. Layard et al. (2008) estimate ρ to be around 1.30 (0.97-1.62) in the UK using the British Household Panel Survey (BHPS) and 1.26 (1.16-1.37) in a combined estimate using BHPS and five other international surveys. This implies marginal utility falls at an even faster rate with respect to income.

The use of the Layard et al. (2008) central estimate, $\rho = 1.3$, is suggested by DWP and HM Treasury (Fujiwara (2010); HM Treasury (2020)).

Calculating weights

In order to calculate welfare weights, the Green Book (HM Treasury (2020)) suggests splitting the net equivalised[4] household income distribution into five quintiles and calculating the median income of each quintile. The weights are then derived as in

[4] Equivalisation adjusts values for greater comparability across households of different sizes and compositions.

equation (5), by expressing the marginal utility of each quintile median as a fraction of the marginal utility of the overall distribution median. Row 2 of table 1 provides an example of the resulting welfare weights, using the HBAI publication (ONS (2020a)). Household-level estimates of net equivalised household income are provided by the ONS annually in their Households Below Average Income (HBAI) publication, the foremost source for National Statistics on household income and inequality in the UK (ONS (2020a)).

The second part of table 1 provides an illustrative CBA to demonstrate the potential power of welfare weighting. Assuming policy makers can target an intervention solely on one quintile of lifetime earners, with a present value earnings benefit of £200m and a cost of £100m. The un-weighted Net Present Value (NPV) is £100m for each option. Hence the standard unweighted CBA metric implies no preference for targeting. By contrast, the weighted NPV favours targeted intervention with the lowest income quintile. Note that the NPV is negative for the highest income quintile: a £100m loss to taxpayers has more utility value than a £200m gain for high-earners.

Table 1: Illustrative example of weighted Net Present Value estimates

	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
Median Earnings £pw*	£256	£392	£514	£685	£1035
Welfare Weight	2.47	1.42	1.00	0.69	0.40
PV Cost	£100m	£100m	£100m	£100m	£100m
PV Benefit	£200m	£200m	£200m	£200m	£200m
NPV (un-weighted)	£100m	£100m	£100m	£100m	£100m
NPV (weighted)	£395m	£184m	£100m	£38m	£-19m

* Median weekly equivalised income by quintile ONS (2020a)

Welfare weights using LEO

Targeted educational policies regularly use measures of parental income to determine eligibility and in theory we could derive weights from parental income. However, the benefit of policies that improve education attainment would likely be accrued by pupils in their later adult life. Thus it is the lifetime earnings of the pupils that is the relevant policy outcome and so the entity that should be used to calculate weighting factors. Parental income, as well as being hard to measure, is not a particularly good proxy for this as it is both an imperfect predictor of child earnings as there is some relative income mobility. This has restrained the use of welfare weights in education policy.

The relatively new Longitudinal Educational Outcomes (LEO) dataset links education administrative data held by DfE with earnings and benefit data from HMRC and DWP. This gives provides the opportunity to robustly estimate lifetime earnings for the first time. Hodge, Little and Weldon (2021), using LEO in combination with the UK Labour Force Survey (LFS), provides a ready made methodology for simulating lifetime earnings

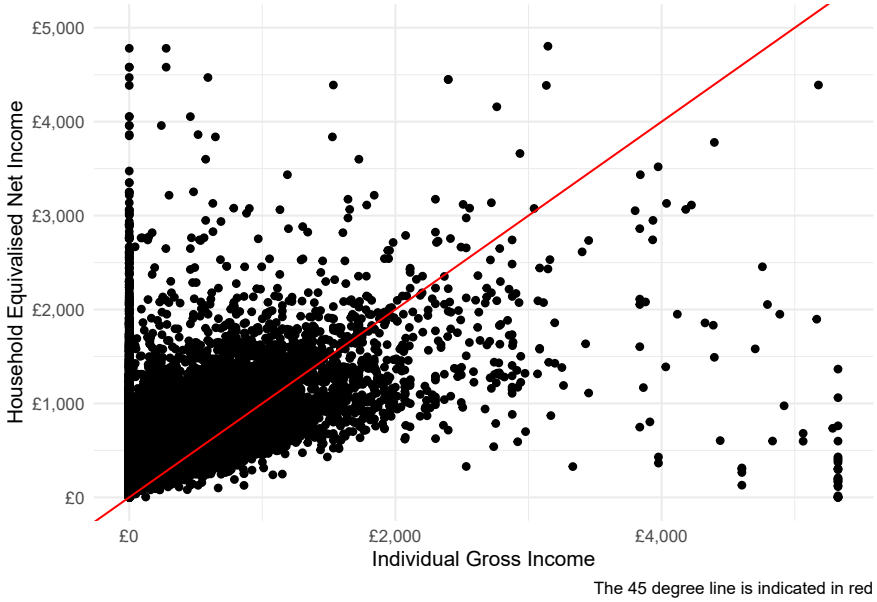
trajectories for pupils. Here we describe some possible approaches for using these lifetime earnings simulations to estimate welfare weights.

The simplest approach is to directly estimate weights using the predicted present values of gross lifetime earnings from Hodge, Little and Weldon (2021). This method does have two potential drawbacks though:

- **The stream of earnings are gross.** We are interested in the utility value of earnings which is better represented by 'money in pocket' and thus earnings net of tax and benefits would be preferred.
- **The stream of earnings are for individuals.** Again, we are interested in the utility derived from earnings, this is better represented by household level earnings and also through adjusting for household composition.

Both these complications are non-trivial to solve as they display complex non-linear relationships. Figure 2 illustrates the relationship between gross individual income and net household equivalised[5] income using data from the Family Resource Survey (FRS (2019)).

Figure 2: FRS: Individual Gross Income vs. Household Equivalised Net Income



A second possible approach is to extract both information on net equivalised household earnings and individual gross income from the Family Resource Survey. This can be used to estimate a fairly 'black box' model that describes the relationship between the two, conditional on a series of control variables. However, age, gender and highest qualification are the extent of the control variables that can be included in any model as

[5]The equivalence values are: first adult (0.67), additional adult (0.33), child aged 14 and over (0.33) and child aged 0-13 (0.2).

we are limited by the fact they need to be present in both the FRS and the LEO data. We explore several different models in the appendix, the preferred model is a Generalised Additive Model (GAM)[6].

A final, unexplored, option would involve simulating both the full tax and benefit system and a model that describes household formation. This would, however, be highly complex, resource intensive and cumbersome to implement.

Estimation results

Table 2 provides a comparison of welfare weights calculated using the income quintiles in the 2018/19 HBAI (ONS (2020a)) with the two methods described above using the LEO/LFS lifetime earnings simulations.

Table 2: Estimated Welfare Weights

	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
HBAI, Net Equiv. Household Earnings	2.47	1.42	1.00	0.69	0.40
LEO/LFS, Individual Lifetime Gross Earnings	3.00	1.46	1.00	0.73	0.49
LEO/LFS, Modelled Net Equiv. Lifetime Household Earnings	1.45	1.16	1.00	0.86	0.69

Note: $\rho = 1.3$

The modelled weights presented in row 3 of table 2 are concentrated more closely to one. This is in part because the distribution of household lifetime income has less variance than the distribution of lifetime individual income but largely this can be attributed to model error, see appendix. It is difficult to improve the modelling approach any further given the data limitations.

We would recommend, where appropriate, either using the HBAI household weights (assuming a perfect rank match of individuals to households) or the individual weights (notwithstanding the two key drawbacks above).

Free School Meals example

Hodge, Little and Weldon (2021) find that changes in GCSE grades have differential associations with lifetime earnings by Free School Meal (FSM) status[7]. Given a policy that

[6] Generalised Additive Models allow for the inclusion of non-parametric or semi-parametric terms, using smoothing functions. This allows an outcome variable to depend linearly on a set of both unknown smooth functions f_i of some predictor variables and more typical linear regression terms, in a structure such as:

$$g(\mathsf{E}(Y)) = \beta_0 + f_1(x_1) + f_2(x_2) + \cdots + f_n(x_n)$$

[7] The return to a one standard deviation improvement in overall GCSE grades for pupils eligible for FSM is estimated to be £82,972 (Female) and £95,839 (Male), compared to £89,214 (Female) and £105,147 (Male) for those pupils not eligible.

has uniform effects across pupils in terms of attainment, in an unweighted CBA an untargeted approach would be preferred to targeting FSM eligible pupils on key value for money metrics.

However, the parental income of FSM eligible pupils will, by definition, be in the bottom quintile of earnings. Due to intergenerational income immobility, FSM eligible pupils will therefore likely fall disproportionately in the lower part of the lifetime earning distribution. Using the lifetime earnings projections from Hodge, Little and Weldon (2021) the median FSM eligible pupil is estimated to fall in the 2nd quintile of lifetime earnings. So a targeted policy would become more favourable under weighting.

Using the full distribution of pupils, the mean average welfare weight for FSM eligible pupils is found to be 1.53 using the HBAI weightings and 1.71 using the individual lifetime earnings projections. Again, these figures should be used with care. The weights are inherently uncertain as discussed and the lifetime earnings projections are based on cohorts of pupils who took their GCSEs between 2002 and 2005.

Issues

The analysis above is not designed to be prescriptive. Whilst we encourage the use of welfare weights, this should be proportionate and for illustrative purposes only. Analysis both with and without applying welfare weights should be presented in any economic appraisal for full transparency.

In addition to welfare weights, one could also consider subjective preferences for redistribution, on equity grounds. The Green Book notes the requirement for Equality Impact Assessments (EIAs), to assess distributional effects based on characteristics other than income. EIAs include (but are not limited to) protected characteristics under the Public Sector Equality Duty. In short, the use of welfare weights is intended to complement, rather than displace, other forms of distributional analyses.

Productivity

Non-technical summary

Skills are of greater value to employers than the wages they pay to their employees. In this section we consider two ways of adjusting GCSE earnings returns, to better represent productivity benefits.

1. **Labour Cost Uplift**: accounts for employer's willingness to pay the 'full' cost of more highly skilled workers. We illustrate the size of this uplift, based on recent ONS data.
2. **Product Market Uplift**: a larger uplift could be applied to account for firm's profit share. We do *NOT* recommend this uplift as there is significant uncertainty over it's level and appropriateness.

Labour cost uplift

Following neoclassical theory, profit-maximising firms employ workers up to point where their marginal cost of labour equals their marginal revenue product. Hence the higher additional costs of employing better-qualified workers can reveal the market rate of those qualifications. Additional costs are inclusive of wage and non-wage labour costs.

The Office for National Statistics publish an Index of Labour Costs per Hour (ILCH), disaggregating *wage costs* and *other costs* (ONS (2020b)).[8] Other costs include estimates of employer National Insurance Contributions, pension contributions, sickness, maternity and paternity payments, and benefits in kind. Note that these benefits largely accrue directly to the worker.

A factor uplift could be applied to earnings, based on the reciprocal share of wage costs (WC) in labour costs (LC). Average Labour Costs were estimated to be £21.20 per hour, of which 85% is attributed to Wage Costs (ONS (2020b)). This would imply a Labour Cost Uplift (LCU) of 1.18, as follows:[9]

$$LCU = \frac{1}{WC/LC} = \frac{1}{0.85} = 1.18 \tag{6}$$

[8]The ILCH data are derived from Monthly Wages and Salaries Survey and the Labour Force Survey. They are published by the ONS as 'experimental statistics'. As such, they are not fully developed and do not yet meet the rigorous quality standards of National Statistics. More information on experimental statistics can be found here.

[9]Index of Labour Costs per Hour, seasonally adjusted, Quarter 1 2020. Estimate for the whole economy, noting that estimates are available by broad sector. Estimates are available here.

Hayward, Hunt and Lord (2014) also estimated productivity returns to GCSEs by applying a 30% uplift to wage returns, to account for non-wage costs.[10] This uplift included various capital overheads, associated with employing more workers (e.g. the cost of office floor space and IT equipment).

Where appraisals consider increases in the level of employment, it may be necessary to adjust for capital costs. The benefits of GCSE attainment are, however, associated largely with remuneration rather than increased employment. Capital overheads represent a semi-fixed (rather than variable) cost with respect to rates of pay. In our view, these overheads ought to be excluded from the uplift in most cases.

Product market uplift

Employers are willing to pay the full cost of labour in both competitive and noncompetitive labour markets. The Labour Cost Uplift is justifiable, irrespective of the degree of competition that one assumes in the labour market.

Where employers have monopsony power they can also generate surplus revenue by employing more productive workers. A Product Market Uplift (PMU) would account for firms' profit share, resulting from public investment in schools to improve skills.

PMUs were considered in Department for Work and Pensions' cost-benefit framework for employment programmes (Fujiwara (2010)). This framework illustrated an uplift of 26%, taking the reciprocal of labour's share of UK factor income (Y).[11]

$$PMU = \frac{1}{LC/Y} = \frac{1}{0.8} = 1.26 \tag{7}$$

DWP noted the following issues with this uplift factor:

1. The extent of monopsony power is uncertain
2. Employees can in some cases be paid more than their marginal product, as a result of:
    (a) Collective bargaining power
    (b) Information asymmetries, which lead employers to overestimate some workers' productivity

[10] The 30% uplift adopted by Hayward, Hunt and Lord (2014) is recommended in research by the Better Regulation Unit in Cabinet Office (2005), principally to account for the cost of regulatory change to business. It includes expenses such as premises, telephone, heating, electricity and IT equipment.

[11] 0.8 is a rounded estimate reported by Fujiwara (2010). It represents the labour share of income relative to others factors of production including entrepreneurship, land and capital. If this uplift was applied in appraisal, one would need to update this estimate.

3. Part of the profit-share accrues to the non-UK owners of UK firms, which are typically excluded in CBA (HM Treasury (2020))

In combination, these issues suggest upward bias in the level of uplift. DWP concluded that an uplift should not be applied in the appraisal of employment programmes. Instead, the appraiser should note that labour costs undervalue total productivity in the CBA. In our view, this is an appropriately conservative position to adopt in the appraisal of education policy.

Issues

Marginal versus average improvements in productivity: the estimates above consider the benefit of a marginal, rather than average, level of improvement in human capital. In most schools policy appraisal, the number of pupils benefiting from the intervention represent a small proportion of the workforce. If a policy significantly raises or lowers the average skills across a large proportion of UK workforce, then it may be more appropriate to estimate average benefits, including a larger estimate of consumer and producer surpluses. If the supply of skills substantively changes, then the analysis would ideally take account of price effects. The treatment of large scale policies relates to the discussion of macroeconomic modeling approaches in the next section. There is a case to treat large impacts differently, making less conservative estimates about the benefits accruing to employers and the wider economy.

Applying productivity uplifts and welfare weights: Productivity uplifts and welfare weights can be applied in the same CBA. Non-wage labour costs mainly benefit the individual worker and are subject to diminishing utility with respect to income. Hence welfare weights can be applied to the labour costs uplift, such as benefits in kind. Whilst employees may not receive these benefits immediately (e.g. employer pension contributions), those at the lower end of the income spectrum are still likely to gain more utility from these benefits than those at the top. One would ideally discount the present value of pensions more heavily than wages, although this may not be practicable. Weights would not be applied to profit shares (the Product Market Uplift) because firm owners are not subject to the same welfare weight.

General and specific skills: There is a judgment to make about the extent to which firms benefit financially from different types of skill. The recommendations above relate generically to primary and secondary education policy. Higher productivity uplifts are often considered in the appraisal of training and vocational programmes (Gambin et al. (2014)). The key distinction is that skills acquired during school are largely transferable between jobs. Transferable human capital is more likely to benefit the worker rather than the firm (Bishop (1994); Blundell et al. (1999)). The theory is that workers can move firms if their

skills are not full rewarded. By contrast, firm-specific training creates productivity gains that can only be captured by the firm (Dearden, Reed and Van Reenen (2006)). One may wish to use different uplifts, for example, where schools invest in vocational skills.

Growth

Non-technical summary

Hodge, Little and Weldon (2021) use a microeconomic model to estimate individual wage returns. By contrast, macroeconomic models can take account of benefits to the whole economy, which often motivate public investment in education.

Leading macro estimates might serve as an 'upper bound' on the total productivity benefit of education (Crawford and Cattan (2013)). These are especially instructive for large scale reforms to the education system. Microeconomic estimates retain the advantage of fine-grained measurement of attainment and the ability to control for pupil and school level characteristics. As such, they are more suited to detailed policy appraisal.

Macroeconomic benefits

DfE commissioned the Institute for Fiscal Studies to review microeconomic and macroeconomic literature on the benefits of education (Crawford and Cattan (2013)). There are several other papers summarising the empirical research linking education and economic growth (for a recent example, see Valero (2021)). We not to repeat these summaries in any detail; rather, we briefly note their practical implications for CBA.

Crawford and Cattan (2013) considered the empirical evidence arising from two types of macroeconomic strategy:

1. **Growth accounting**, whereby human capital is incorporated as a factor input in the production function.
2. **Macro growth regressions**, aiming to capture the dynamic interactions between educational performance and the whole economy. This recognises that returns to schooling are endogenous (Card (2001)).

Hanushek and Woessmann have advanced the argument that long-run growth is overwhelmingly caused by "knowledge capital" (Hanushek and Woessmann (2015)). Their macro growth regressions have been among the most prominent in education policy for many years, particularly within the European Union and the OECD. Their estimates explore cross-country variation in growth, education and other factors at national level. The consequences of higher cognitive skills reportedly represent "multiples of GDP" (Hanushek and Woessmann (2015)). Recent contributions have considered:

1. **Positive scenarios**, based on the European Union's educational goals (Hanushek and Woessmann (2020))
2. **Negative scenarios**, based on the potential learning losses resulting from school closures during the Coronavirus pandemic (Hanushek and Woessmann (2020))

One of Hanushek and Woessmann's positive scenarios sees an average student achievement increase by 25 PISA points. Over the next 80 years, they estimate that this would add €9.7 trillion in present value to UK GDP, over the status quo. That was equivalent to 340% of current GDP at the time of estimation, in early 2020. It represents an increase in the long-run growth rate of 0.5% (Hanushek (2020)).[12]

Their estimates of Coronavirus-induced learning loss are also sizable. For instance, one third of a year of lost learning is estimated to reduce UK GDP by over $2 trillion over the next 80 years.[13]

It is difficult to compare Hanushek and Woessmann's estimates directly to the wage returns reported by Hodge, Little and Weldon (2021). Rough comparisons would suggest that macroeconomic estimates are multiple times higher.

Issues

Whilst evidence for sizable growth benefits is well-founded, in practice this is difficult to apply in policy appraisal. Crawford and Cattan (2013) concluded that the reliance on national level data means that:

> "…[growth regressions] may not produce an estimate of the true causal effect of education on economic growth; instead this approach is likely to produce an upper bound"

Crawford and Cattan (2013) provide a detailed account of these methodological issues, covering measurement error, endogeneity and model specification. One such issue is that Hanushek and Woessmann pool data from various countries, so their estimates represent

[12] Discounted value of future increases in GDP until 2100 due to the reform, expressed in billion Euro (PPP), as a percentage of current GDP, and as a percentage of discounted future GDP. "GDP increase in year 2100" indicates by how much GDP in 2100 is higher due to the reform (in %). "Long-run growth increase" refers to increase in annual growth rate (in percentage points) once the whole labor force has reached higher level of educational achievement. "Increase in PISA score" refers to the ultimate increase in educational achievement due to the reform. See text for reform parameters.

[13] Estimate for the UK is 2,154 billion US Dollars. The baseline is GDP for 2019, in billions of USD and in 2017 purchasing power parity (PPP) terms from the World Bank. Present value of lost GDP is based on estimated difference in GDP for 80 years with lower achieving labour force expected from educational losses of one-third or two-thirds years compared to future GDP without learning loss. Future losses are discounted at 3 percent.

the average growth effect across these nations. The average effects are then applied to estimate each country's GDP growth scenarios. This is an imperfect guide for UK policy makers as the country's relationship between education and growth could differ markedly from the average.

Microeconomic estimates allow for more detailed, policy-level analysis. It is possible to control for a range of pupil-level factors, which might explain the differences in educational performance and later outcomes. For instance, if we wish to assess a policy intervention in secondary school, then it is important to control for prior attainment in primary school. If we wish to target policy intervention on a specific sub-group, then we would ideally assess the returns associated specifically with that group.

Causal interpretation remains elusive in most macro- and micro-estimation strategies and so this is a matter of degree: *how far can either strategy control for confounders and get closer to the causal link between education policy and increases in economic output?*

Microeconomic estimates are generally preferred in CBA and provide more conservative the scale of the policy benefits. An appraisal could, in this case, usefully reference evidence of the macroeconomic benefits. For large-scale interventions, affecting a large proportion of the future labour force, it may be useful to approximate the long-term growth effects within the CBA.

Signalling

Non-technical summary

The well-established link between education and later earnings does not necessarily mean that education makes individuals more productive. GCSE grades, for example, could either reflect skills acquired in school, or signal pupils' pre-existing abilities.

We do not consider that signalling effects undermine the use of earnings returns, such as those found in Hodge, Little and Weldon (2021). The value of school attainment as a rough measure of individual skill has been verified by a wide variety of studies of labor market outcomes (see Card (2001); Wyness, Macmillan and Anders (2021)). Nevertheless, it is important to consider how relevant signalling might be for a given policy.

Signalling effects are context-specific. The factors to consider include; subject, grade distribution, pupil cohort, time period, later educational pathways, occupations and industries.

Theory and evidence

Following human capital theory, higher earnings could result from the productivity-enhancing effects of compulsory schooling. On the other hand, test scores could signal a learner's innate motivation and abilities, irrespective of whether schooling augments productivity.[14]

Gambin et al. (2014) argue that it "makes little difference" to the individual learner whether higher earnings arise from the effects of human capital or signalling. Yet the distinction is important to policy makers. Human capital theory provides a more compelling case for public investment to education in order to grow the economy as a whole. In the presence of signalling, lifetime earnings could overestimate the return on human capital investment.

The relative importance of human capital and signalling is an empirical question, for which there is a wide literature, but few studies focusing specifically on GCSEs.

Machin, McNally and Ruiz-Valenzuela (2020) report significant returns to gaining a 'C' grade in English GCSE on appeal, rather than a 'D'. They describe this as a 'high stakes exam', in part because achieving at least a 'C' grade is used by employers and higher educational institutions, to screen applicants. Given these pupils fall (with some randomness) marginally either side of the 'C' grade cut-off, they might be expected to

[14] Human capital theory follows in the tradition of Mincer (1958) and Becker (1960); (1964). Signalling theory was developed by Spence (1973) among others.

have very similar levels of human capital. This could be taken as evidence of a signalling effect. We cannot generalise from this finding though, as the effect may be very different for other subjects or at other grade boundaries.

It is important to recognise potential signalling effects in CBA but we do not believe there is sufficient concern to invalidate the use of earnings returns, such as those estimated in Hodge, Little and Weldon (2021). The value of school attainment, as a rough measure of individual skill, has been verified by a wide variety of studies of labor market outcomes. David Card (2001) is among several leading economists to empirically measure human capital effects. He reviews studies that use a range of robust experimental methods and instrumental variables for completed school education. He concludes that:

> "the returns to schooling are typically as big or bigger than the corresponding ordinary least squares estimates"

In a more recent briefing note, Wyness, Macmillan and Anders (2021) concluded that

> "The most convincing quantitative studies from the literature suggests that signalling plays a relatively limited role. This, coupled with causal evidence of the wider non-pecuniary benefits of education, implies that failing to invest in education, particularly at critical ages and stages, would be a very risky strategy for governments to adopt."

In our view, this consensus ought to be reflected in our general approach to CBA for school-based policies. On balance, making a downward adjustment to estimates of the private earnings returns, to account for signalling effects, is more likely undervalue productivity benefits, given that these are already somewhat conservative.

Issues

Identification:: A cautious conclusion is that both human capital and signalling effects are evident, with no consensus on their relative importance with respect to Key Stage 4 exams. In estimating returns to GCSE grade improvements, Hodge, Little and Weldon (2021) do not use experimental or instrumental variable techniques, because these methods are not available across all grades, subjects and pupil sub-groups. We require estimates to inform a wide range of education policy appraisals. The pursuit of robust causal identification inevitably trades-off against the need for breadth of policy application. Given the difficulties in identifying causal impacts within the model, we cannot discount signalling effects; equally, it may be underestimating human capital effects.

Policy context: The literature suggests that signalling effects are context-specific. Their effect may vary by GCSE subject, grade distribution, pupil cohort, time period, later educational pathways, occupations and industries. Human capital effects would also be expected to vary over time, due to changes in the curriculum or teaching quality, or due to the value of skills in the future labour market.

We do not recommend a routine downward adjustment for signalling, noting that there is an equally strong case for an upward adjustment to better capture human capital effects. This can only be treated as a general rule and concern for signalling effects ought to be addressed in each policy appraisal.

Wellbeing

Non-technical summary

Traditional economic indicators (e.g. income and GDP) cannot fully capture the effects of schools on people's lives. Giving regard to direct measures of personal wellbeing can improve decision-making.

Wellbeing research can be used at all stages in the policy-making cycle, from formulating options through to evaluation. At the 'long-list' appraisal stage, wellbeing impacts can be included among the Critical Success Factors (CSFs) for a project. At short-list stage, we need a common currency to capture productivity and wellbeing benefits in the same analysis. There are two leading options:

1. Value wellbeing outcomes in monetary terms
2. Value productivity outcomes in units of wellbeing

The first option is an extensive form of CBA. This method has been advanced internationally, notably in New Zealand and Australia (NZ Treasury (2015), (2019); Australian Social Value Bank (no date)).

The second option can take the form of a Cost Effectiveness Analysis (CEA). In CEA, costs are in one set of units (money) and the benefits in another (e.g. happiness-years). CEA is already well-established in the appraisal of health expenditure. The UK Government evaluates gains in health using Quality-Adjusted Life-Years (QALYs). Wellbeing CEA proposes a different unifying measure of benefit, based on subjective wellbeing.

Wellbeing appraisal continues to evolve in the UK, and there are some data and evidence constraints in its application to schools policy. We could not, for instance, emulate our method to predict wage impacts of Key Stage 4 performance (Hodge, Little and Weldon (2021)), to make similar predictions of adult life satisfaction.

These limitations relate largely to appraisal, i.e. before a policy is implemented, where we are reliant on estimates from the existing evidence base. Limitations can be overcome at the evaluation stage, i.e. after the policy is implemented, as there are proven and reliable methods of evaluating wellbeing for children and young people.

Impact of schools on wellbeing

Outcomes at age 16 have a significant bearing on wellbeing throughout the life course. Using data from two British cohort studies, Frijters, Johnston and Shields (2013) reported that variables observed up to age 16 predicted around 7% of the variation in average adult life satisfaction. Adding the contemporaneous effects of childhood on adulthood variables increases that predictive power to 15.6%, while adding long lags of life satisfaction increased it further still, to 35.5%.

The consensus is that life satisfaction can be improved within the education system (Longhi et al. (2018)). For example, Layard et al. (2018) estimated the impact that schools have on three different outcomes at age 16 based on measures of emotional health, conduct and GCSE test scores. They used data from the Avon study, which also allowed controls for family characteristics.[15] The fixed effects of both primary and secondary school attended were sizable, across all three outcomes (see table 3).

Table 3: How children's outcomes at age 16 are affected by school and family (Standardised coefficients)

	Emotional health at 16	Behaviour at 16	GCSE score at 16
Primary School	0.27	0.32	0.21
Secondary school	0.28	0.31	0.38
Family income (log, averaged)	0.07	0.08	0.14
Parents' education (years)	–	0.04	0.17
Father unemployed (% of years)	–	–	-0.03
Mother worked (% of 1st year)	–	–	-0.02
Mother worked (% of other years)	–	-0.05	0.04
Parents' involvement with child	0.04	0.05	0.02
Parents' aggression to child	-0.03	-0.12	–
Mother's mental health	0.16	0.17	0.03
Father's mental health	0.04	–	–
Conflict between parents	-0.04	-0.14	-0.01

Source:
Avon Longitudinal Study of Parents and Children (ALSPAC).
Cross-section analysis reported in table 16.4 of Layard et al. (2018).
Available at http://cep.lse.ac.uk/origins/onlinematerial.pdf.

Primary schools appear to have as much influence on behaviour and emotional health as secondary schools, even though these measures are taken at age 16. The same study reports that individual primary school teachers had even larger impacts on their children's emotional health than on their learning of maths. The effects of primary school teachers were detected up to 10 years later (Layard et al. (2018)).

[15] The data is from the Avon Longitudinal Study of Parents and Children (ALSPAC). The survey covered around 70% of all children born in the Bristol area, between April 1991 and December 1992. School effects could be estimated using dummy variables for each school. The authors do not refer to these as 'fixed effects' but the method is broadly equivalent. The advantage of the dummy variables is that they can detect the whole-school effect. The drawback, from a policy perspective, is that we cannot observe the specific characteristics of schools that might be associated with better outcomes.

The combined effect of family variables included in the model is similar in magnitude to the effect of secondary schools.[16] The single largest family determinant of a child's wellbeing and behaviour is the mental health of the mother. The biggest family factors affecting academic performance were family income and parents' education.

Several studies have estimated the impact of further study, proxied by years of education or the highest qualification level achieved. The effects are mixed, weak and in some cases negative. Where effects are present, they are partly mediated through other channels such as income and reduced criminality.

Overall, Layard et al. (2018) concluded that:

> "If we wish to predict which children will lead satisfying adult lives, the best indicator is their emotional health at age 16. This is more important than their academic qualifications right up to the age of 25 – and more important than their behaviour in childhood."

From an appraisal standpoint, practitioners would need to consider whether the policy intent is to improve school or teacher *quality* (with strong effects on child and adult wellbeing) or to increase the *quantity* of education by encouraging further study (with mixed effects on wellbeing).

Wellbeing appraisal

The next question is how practitioners might best take account of these wellbeing impacts in school policy appraisal.

Wellbeing evidence can be used from the outset, to determine the policy goals. This is already the case in a number of areas. For example, schools have been testing a new curriculum, developed as part of a programme to teach resilience. It includes lessons in areas that are important for wellbeing, including; relationships, healthy habits, social media awareness, and mindfulness (Bounce Forward (no date)).

Wellbeing objectives can then be reflected in the Critical Success Factors (CSF) for a project, by which the 'long-list' of options are appraised (see HM Treasury (2020)). For example, Defra and the Cabinet Office developed a framework for Multi-Criteria Analysis (MCA) incorporating wellbeing outcomes (Maxwell et al. (2011)).

[16] The inclusion of school dummies was possible because there are multiple pupils per school in the sample. By contrast, there was only one child in each family. The model could not include dummy variables to estimate the overall effect of the child's family. Hence the comparison - based on adding the effects of observable family characteristics - will miss any unobservable family effects.

At short-list stage, CBA has been the primary device in government appraisal. In this case, all social costs and benefits are translated into monetary terms. A number of studies have considered how best to evaluate wellbeing in money terms. In education, monetary values have been placed on the wellbeing impacts of adult learning (Dolan and Fujiwara (2012)), and for higher education outcomes in evaluations of the National Citizens Service (Jump and Simetrica (2017); Dokal et al. (2020)).

The method of converting people's preferences and utility into monetary figures can create systemic errors (Bronsteen, Buccafusco and Masur (2013)). These issues have led some economists to call for alternative appraisal tools, described generically as wellbeing cost-effectiveness analysis.[17] In this format, costs are in one set of units (money) and the benefits in another (e.g. cumulative 'happiness-years').[18]

CEA is already used in government for health expenditure (see Glover and Henderson (2010); NICE (2013)). For all the possible treatments, the Government's economic guidelines are to evaluate health gains in units of Quality-Adjusted Life-Years (QALYs). Hence the appraisal method is well-established as a key decision-making tool. The difference in wellbeing CEA is the change units, from quality of life to subjective wellbeing. Happiness-years may be more instructive in education policy than QALYs, which focus largely (though not exclusively) on physical health.

The Green Book legitimises both CBA and wellbeing CEA, with the choice governed by the nature of the policy and available evidence. The Green Book suggests that using wellbeing as the primary outcome variable:

> "may be particularly useful in certain policy areas, for example community cohesion, children and families" (p42, HM Treasury (2020))

Where the measurements of the policy are mainly money terms (e.g. wage impacts) then it may be natural to stick to those units throughout the analysis (O'Donnell et al. (2014); HM Treasury (2020)).

Issues

An achievable aim in this section of the Handbook is to highlight the relevance and potential applications of wellbeing economics, in appraisal.

Data that track people's actual experience of life are consistently found to be reliable and valid for policy appraisal and research purposes. There are, however, some issues that would need to be explored within each policy appraisal.

[17] See, for example, the All Party Parliamentary Group on Wellbeing Economics' Open Letter to the Chancellor, *A Spending Review to Increase Wellbeing* (Layard (2019)).

[18] A free video course on these methods is available on the What Works Centre for wellbeing website here.

Data availability: we need to consider the long period over which educational outcomes materialise. The estimates in Hodge, Little and Weldon (2021) are reliant on large scale administrative data, containing detailed information on GCSE performance and incomes up to 14 years later. We do not have the same quality of data to assess later wellbeing impacts, at present.

Wellbeing analysis 'at the margin': Wellbeing impacts can be easier to evaluate where there is a clear change of state, such as flooding events, lottery wins and redundancy. This is evident in the wellbeing literature, where education tends to be measured in level shifts, such as years of schooling or highest qualification. In most school policy appraisal, we need to measure marginal changes in attainment, such as GCSE grades. Incremental improvements in attainment are less likely to register on subjective wellbeing measures, again because we lack the large administrative data to identify those changes, robustly.

Valuation: There is no consensus on the monetary values one might place on life satisfaction of children. Wellbeing appraisal remains viable, but the valuations would need to be selected carefully and subject to sensitivity analyses.

A combination of these issues can be more acute before policy implementation, where the appraisal relies upon proxy values from existing wellbeing research. They can be overcome through high quality evaluation plans, which have the potential to enrich the existing literature.

Glossary

Appraisal: We use this term to refer to the process of assessing the costs, benefits and risks of a policy. We draw a distinction between appraisal (the assessment process *before* a policy is implemented) and evaluation (*after* implementation) although these are sometimes used interchangeably.

CBA: Cost Benefit Analysis. We use CBA as a shorthand for *Social* Cost Benefit Analysis, which aims to assess the impact of policy intervention on social welfare. In CBA outcomes are valued in monetary terms, unless it is not proportionate or possible to do so.

CEA: Social Cost-Effectiveness Analysis is a variant of Social CBA. It compares the costs of alternative ways of producing the same or similar outputs.

DfE: Department for Education

DWP: Department for Work and Pensions

GCSE: General Certificate of Secondary Education

Green Book: Central Government guidance on appraisal and evaluation. See HM Treasury (2020).

KS4: Key Stage 4 (KS4) is the legal term for the two years of school education which incorporate GCSEs, and other examinations, in maintained schools in England normally known as Year 10 and Year 11, when pupils are aged between 14 and 15 by August 31st.

OECD: Organisation for Economic Co-operation and Development

PISA: PISA is the OECD's Programme for International Student Assessment. PISA measures 15-year-olds' ability to use their reading, mathematics and science knowledge and skills to meet real-life challenges.

QALY: Quality-adjusted life-year, a measure of disease burden including both the quality and the quantity of life lived. One QALY equates to one year in perfect health.

Appendix

Welfare Weights Modelling

Model specifications

We compare the performance of 3 different models, all using fully interacted covariates of age, sex and highest qualification obtained:

- Linear Regression
- Poisson Regression
- Generalised Additive Model (GAM)[19]

	Adj. R^2	RMSE
Linear	0.33	0.66
Poisson	0.28	0.68
GAM	0.36	0.64

The GAM provides the best fit to the data; it has the highest adjusted R^2 and the lowest RMSE.

Model diagnostics

We estimate a set of welfare weights by estimating household equivalised net earnings using the GAM model on the FRS data. If the model was highly accurate the resulting weights would approximately equal the weights derived from the HBAI. The two sets of weights are compared in table 4. The modelled weights are much closer to one, this is intuitive given the poorer predictive power at the tails of the distribution, as seen in the Q-Q plot in figure 3.

[19] Index of Labour Costs per Hour, seasonally adjusted, Quarter 1 2020. Estimate for the whole economy, noting that estimates are available by broad sector. Estimates are available here.

Figure 3: Q-Q and Residual Diagnostic Plots

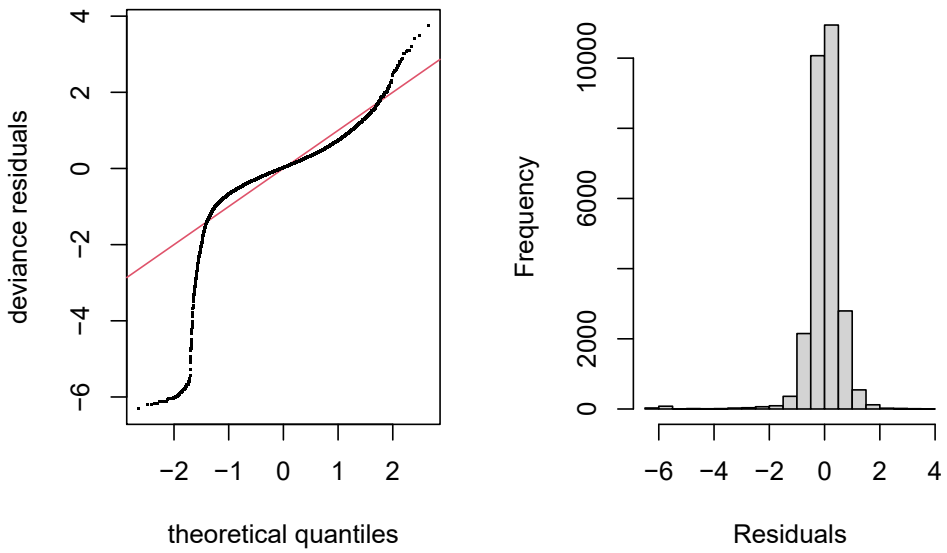


Table 4: FRS Welfare Weights - Actual vs. Modelled

	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
Actual	2.47	1.42	1.00	0.69	0.40
Modelled	1.61	1.22	1.00	0.75	0.50

Note: $\rho = 1.3$

References

Australian Social Value Bank (no date) *Australian Social Value Bank - makes measuring your social impact simple!* Australian Social Value Bank.

Becker, G. (1960) 'Underinvestment in college education?', *American Economic Review*, 50, pp. 346–354.

Becker, G. (1964) *Human capital: A theoretical and empirical analysis, with special reference to education*. National Bureau of Economic Research.

Better Regulation Executive (2005) *Measuring Administrative Costs: UK Standard Cost Model Manual*. Cabinet Office.

Bishop, J. (1994) 'The Impact of Previous Training on Productivity and Wages', in Lynch, L. (ed.) *Training and the Private Sector*. University of Chicago Press, pp. 161–200.

Blundell, R. L., Dearden, L., Meghir, C. and Sianesi, B. (1999) 'Human Capital Investment: The Returns from Education and Training to the Individual, the Firm and the Economy.', *Fiscal Studies*, 20(1), pp. 1–23.

Bounce Forward (no date) *Healthy Minds research project*. Bounce Forward.

Box, G. E. P. (1976) 'Science and statistics', *Journal of the American Statistical Association*, 71, pp. 791–799.

Box, G. E. P., Luceño, A. and del Carmen Paniagua-Quinones, M. (2009) *Statistical Control by Monitoring and Adjustment*. Wiley.

Bronsteen, J., Buccafusco, J. and Masur, J. S. (2013) 'Well-being analysis vs. Cost-benefit analysis', *Duke Law Journal*, 62, pp. 1603–1689.

Card, D. (2001) 'Estimating the return to schooling: Progress on some persistent econometric problems.', *Econometrica*, 69(5), pp. 1127–1160.

Clark, A. and Oswald, A. J. (2002) 'A simple statistical method for measuring how life events affect happiness.', *International Journal of Epidemiology*, 31, pp. 1139–1144.

Crawford, C. and Cattan, S. (2013) *Assessing the economic benefits of education: Reconciling microeconomic and macroeconomic approaches*. CAYT Report No.4. Centre for Analysis of Youth Transitions.

Cunha, F. and Heckman, J. (2008) ‘Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation.’, *Journal of human resources*, 43(4), pp. 738–782.

Cunha, F., Heckman, J. and Schennach, S. (2010) ‘Estimating the Technology of Cognitive and Noncognitive Skill Formation.’, *Econometrica*, 78(3), pp. 883–931.

Dearden, L., Reed, H. and Van Reenen, J. (2006) ‘The Impact of Training on Productivity and Wages: Evidence from British Panel Data’, *Oxford Bulletin of Economics and Statistics*, 68(4), pp. 397–421.

Department for Work and Pensions, Office for National Statistics and NatCen Social Research (2019) ‘*Family resources survey, 2017-2018*’. UK Data Service.

Di Tella, R. and MacCulloch, R. (2010) ‘Happiness adaptation to income and to status in an individual panel’, *Journal of Economic Behavior & Organization*, 76, pp. 834–852.

Diener, E. and Lucas, R. (1999) ‘Personality and subjective wellbeing’, in Kahneman, D., Diener, E., and Schwarz, N. (eds) *Wellbeing: The foundations of a hedonic psychology*. Russell Sage Foundation, pp. 213–229.

Dokal, B., McGee, A., McKay, E., Fitzpatrick, A., Hodgson, E., Matthews, P., Nickson, T., Bates, J., Miller, J. and Conlon, G. (2020) *National Citizen Service 2018 Evaluation: Main report*. DCMS Research Report. Department for Digital, Culture, Media & Sport.

Dolan, P. and Fujiwara, D. (2012) *Valuing Adult Learning: Comparing Wellbeing Valuation to Contingent Valuation*. BIS Research Paper No 85. Department for Business Innovation & Skills.

Dreze, J. and Stern, N. (1987) ‘The Theory of Cost-Benefit Analysis’, in Auerbach, A. J. and Feldstein, M. (eds) *Handbook of Public Economics, vol. 2*. Elsevier, pp. 909–989.

Easterlin, R. (1974) ‘Does economic growth improve the human lot? Some empirical evidence’, in David, R. and Reder, R. (eds) *Nations and households in economic growth: Essays in honor of Moses Abramovitz*. Academic Press, pp. 89–125.

Frijters, P., Johnston, D. W. and Shields, M. A. (2013) ‘Does childhood predict adult life satisfaction? Evidence from British Cohort Surveys’, *The Economic Journal*, 124, pp. F688–F719.

Fujiwara, D. (2010) *The Department for Work and Pensions Social Cost Benefit Analysis framework*. Working Paper No 86. Department for Work and Pensions.

Gambin, L., Beaven, R., Hogarth, T., May-Gillings, M. and Long, K. (2014) *Methodological Issues in Estimating the Value Added of Further Education, Higher Education and Skills: A review of relevant literature*. BIS Research Paper No. 166. Department for Business Innovation & Skills.

Glover, D. and Henderson, J. (2010) *Quantifying health impacts of government policies: A how-to guide to quantifying the health impacts of government policies*. Department of Health.

Gregg, P., Jonsson, J., Macmillan, L. and Mood, C. (2017) 'The Role of Education for Intergenerational Income Mobility: A comparison of the United States, Great Britain, and Sweden', *Social Forces*, 96(1), pp. 121–152.

Hanushek, E. A. and Woessmann, L. (2015) *The Knowledge Capital of Nations: Education and the Economics of Growth*. Cambridge, MA: MIT Press Pages.

Hanushek, E. A. and Woessmann, L. (2020) 'A quantitative look at the economic impact of the European Union's educational goals', *Education Economics*, 28(3), pp. 225–244.

Hanushek, E. and Woessmann, L. (2020) *The economic impacts of learning losses*. OECD Education Working Papers 225. Organisation for Economic Co-operation and Development.

Hayward, H., Hunt, E. and Lord, A. (2014) *The economic value of key intermediate qualifications: Estimating the returns and lifetime productivity gains to GCSEs, A levels and apprenticeships*. Research Report. Department for Education.

HM Treasury (2020) *The Green Book: Central government guidance on appraisal and evaluation*.

Hodge, L., Little, A. and Weldon, M. (2021) *GCSE Attainment and Lifetime Earnings*. Research Report. Department for Education.

Jump and Simetrica (2017) *If you could bottle it... A wellbeing and human capital value for money analysis of the NCS 2015 programme*.

Keeley, B. (2007) *Human Capital: How what you know shapes your life*. OECD Insights. OECD Publishing.

Layard, R. (2019) *A Spending Review to Increase Wellbeing: : An open letter to the Chancellor, 2019*. All Party Parliamentary Group on Wellbeing Economics.

Layard, R., Clark, A., Flèche, S., Ward, G. and Powdthavee, N. (2018) *The origins of happiness: The science of wellbeing over the life course.* Princeton, NJ: Princeton University Press.

Layard, R., Nickell, S. and Mayraz, G. (2008) 'The marginal utility of income.', *Journal of Public Economics,* 92, pp. 1846–1857.

Longhi, S., Nandi, A., Bryan, M., Connolly, S. and Gedliki, C. (2018) *Education to employment: The role of wellbeing*. What Works Centre for Wellbeing.

Machin, S., McNally, S. and Ruiz-Valenzuela, J. (2020) 'Entry through the narrow door: The costs of just failing high stakes exams', *Journal of Public Economics*, 190, pp. 104–224.

Maxwell, S., Henderson, D., McCloy, R. and Harper, G. (2011) *Social Impacts and Wellbeing: Multi-criteria analysis techniques for integrating nonmonetary evidence in valuation and appraisal. A discussion of current approaches and opportunities*. Defra Evidence and Analysis Series Paper 5. Department for Environment, Food & Rural Affairs.

Mincer, J. (1958) 'Investment in Human Capital and Personal Income Distribution', *Journal of Political Economy*, 66(4), pp. 281–302.

NICE (2013) *Guide to the methods of technology appraisal 2013*. National Institute for Health; Care Excellence (NICE).

NZ Treasury (2015) *Guide to Social Cost Benefit Analysis*. New Zealand Government Treasury.

NZ Treasury (2019) *The Treasury's CBAx Tool*. New Zealand Government Treasury.

O'Donnell, G., Deaton, A., Durand, M., Halpern, D. and Layard, R. (2014) *Wellbeing and Policy*. Legatum Institute.

ONS (2020a) *Households below average income: 1994/95 to 2018/19*.

ONS (2020b) *Index of Labour Costs per Hour, seasonally adjusted*.

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

Reynolds, A., Ou, S., Mondi, C. F. and Hayakawa, M. (2017) ‘Processes of Early Childhood Interventions to Adult Well-Being’, *Child Development*, 88(2), pp. 378–387.

Spence, M. (1973) ‘Job Market Signaling’, *The Quarterly Journal of Economics*, 87(3), pp. 355–374.

Valero, A. (2021) *Education and economic growth*. Programme on Innovation and Diffusion Working Paper 006. London School of Economics.

Wyness, G., Macmillan, L. and Anders, J. (2021) *Does education raise people’s productivity or does it just signal their existing ability?* CEPEO Briefing Note Series 12. Centre for Education Policy and Equalising Opportunities (CEPEO), UCL Institute of Education.

Department for Education

Reference: DFE-RR1132

ISBN: 978-1-83870-265-6

For any enquiries regarding this publication, contact us at:
Allan.LITTLE@education.gov.uk or www.education.gov.uk/contactus

This document is available for download at www.gov.uk/government/publications