

Early literacy matters: Economic impact and regional disparities in England

Karol Rodriguez Cabrera, Rachel Gomez and Jon Franklin

February 2024



Pro Bono Economics uses economics to empower the social sector and to increase wellbeing across the UK. We combine project work for individual charities and social enterprises with policy research that can drive systemic change. Working with 900 volunteer economists, we have supported over 500 charities since our inception in 2009.

The National Literacy Trust is an independent charity that empowers children, young people, and adults with the literacy skills they need to succeed. Reading, writing, speaking and listening skills give you the tools to get the most out of life, and the power to shape your future. From first words, through school days to training, jobs and beyond, together they're helping people change their stories.

> KPMG LLP, a UK limited liability partnership, operates from 20 offices across the UK with approximately 18,000 partners and staff. KPMG is a global organization of independent professional services firms providing Audit, Legal, Tax and Advisory services. For over 150 years, we've been supporting businesses to grow, our people to achieve and our communities to thrive.



Contents

Summary	4
Introduction	7
PBE's approach	13
Results of the analysis	22
Implications of the analysis	28
Conclusion	30
Annexes	31



Summary

A substantial number of children are being let down from an early stage and are not reaching their potential in communication, language, and literacy (CLL) skills. This not only impairs their development and prospects but costs the economy too.

The early years are the most critical time for the development of a child's abilities to master language, to read and to communicate. However, with early years settings under pressure and families not receiving adequate support, more than a quarter of five-year-olds in England - 187,000 children – did not meet the expected standard for literacy in 2022/23.

It is likely that many of these children could reasonably have achieved the expected level in literacy if they were provided with more tailored support. The current early years standards were only introduced in 2021/22. This means there is not yet robust evidence on the long-term impact of these children not reaching the expected level for the current standards. However, by linking previous versions of early years literacy assessments to later economic outcomes, it is possible to estimate that an additional 106,000 children could reasonably have achieved the expected level in literacy in 2018/19 if they had received the additional support they needed.

The economic cost of insufficient early literacy support is likely to be significant. Each year group of children who do not meet the expected early years standard generates lifetime economic costs of around £830 million. This equates to £7,800 for each five-year-old who could reasonably have been expected to meet the standard. Much of this loss stems from the knock-on effects on academic success and later employment outcomes. The typical child who does not meet the literacy standard at age five loses out on around £5,300 of earnings over their lifetime and costs the government around £2,500 more over their lifetimes through higher spending on education and welfare, as well as lower tax revenue.

These costs can be particularly important in more deprived communities. The connections between deprivation, literacy skills and long-term employment outcomes can create a cycle of disadvantage – the process by which the life outcomes of people born into disadvantage are worse, which impacts their children's life outcomes too. For example, the long-term cost of children that do not meet the standard for literacy at age five in more deprived city areas such Birmingham, Liverpool and Manchester is approximately a combined £30 million for each year group of children. But the effects will reverberate within these areas for generations to come.

As a result, solutions to this challenge need to take into account deprivation, and its relationship to weaker home learning environments. There are many

reasons for this, with lower income caregivers often possessing poorer literacy skills themselves, having fewer financial resources to spend on ageappropriate books, toys and days out, having less time, living in areas with fewer amenities, and having fewer role models – to name just a few. This weaker home learning environment means that solutions to low literacy among young children should consider the role of communities in supporting early childhood education and care settings.

Community solutions to low early literacy need to combine the best of the public, private and social sectors to make a sustained positive difference. Early years settings can bring expertise and structure, social sector organisations can reach out into the community, and employers can bring their resources to bear to unleash the potential of the next generation of learners.

The 'Early Words Matter' campaign from the National Literacy Trust provides an example of such a solution. This campaign targets two-fifths of the most deprived local authorities in the UK with some of the highest numbers of children not meeting the standard for literacy at age five. The National Literacy Trust takes a localised approach, engaging early years settings, employers and local community groups to reach parents and carers and empower them to support their child's early CLL development. If such an approach is successful, the benefits could last for generations by breaking the cycle of disadvantage.

While the challenge of improving CLL skills for children born now and in the future is substantial, differences in outcomes between regions and local authorities provide hope that it is achievable. We can make a big difference by building on the lessons of what has worked in those areas that are performing well and using this to further shape evidence-based, tailored interventions in areas where children are struggling the most. There are potentially huge gains in creating a more even playing field; the number of children not meeting the expected standard for literacy could be more than halved if the worst performing areas in England were supported to reach the standards of the best performing areas.

Seizing that challenge now is critical. Improvements in children's literacy performance in England had levelled off prior to 2020. The situation has since been worsened by the pandemic as lockdown limited access to education and care in formal early years settings. The cost-of-living crisis is also likely to have exacerbated this issue, as it has further reduced parents' ability to support children's learning at home and has led to fewer parents engaging in home learning activities with their children. Effective community interventions that target early literacy support to children in the most deprived areas of England are urgently needed.

106,000

additional five-year-old children could have been meeting the expected literacy standard each year with the right support

Failing to support children to reach standard levels of literacy could result in lifetime costs to society of

£830m

for each cohort of five-yearolds

> The lifetime cost of children that do not meet the standard for literacy at age five in Birmingham, Liverpool and Manchester is a combined

£30m

for each cohort of five-yearolds

If every region was in line with the best performers, it would **more than halve**

the number of children not reaching the expected standard in early years literacy

Introduction

Supporting young children to develop their communication, language and literacy (CLL) skills benefits them and society. The development of reading, writing, speaking, listening and comprehension skills are crucial as they lay the foundations for education, enabling children to engage with and understand new concepts, expand their knowledge, enhance their memory and concentration skills.¹ Additionally, the Department for Education (DfE) has identified that elevating early literacy and numeracy levels can significantly improve employment outcomes and lifetime earnings.²

Unfortunately, there is considerable evidence that suggests hundreds of thousands of children in England are not reaching their full potential in CLL skills each year. Results from the Early Years Foundation Stage (EYFS), the statutory review to assess five-year-olds development, found literacy was the area in most urgent need of improvement, with 187,000 five-year-olds (30%) not achieving the expected level of literacy in 2022/23. Also, 125,000 five-year-olds - 20% of the cohort – did not reach the expected level in communication and language.³

These numbers were lower before the pandemic. In 2018/19, 27% of children did not meet the standard for early years literacy and 18% did not meet the standard for communication and language. While it is important to note that the EYFS was reviewed significantly in 2021, data published before this review suggests that there was good progress in helping children reach the expected CLL levels, although this progress slowed and levelled off in the years prior to the pandemic, as displayed in Figure 1.

¹ Bree & Zee, <u>The unique role of verbal memory, vocabulary, concentration and self-efficacy in children's</u> <u>listening comprehension in upper elementary grades</u>, 2021.

² Department for Education, <u>The economic benefits of effective Reception classes in England</u>, January 2023.

³ Department for Education, <u>Early Years Foundation Stage profile results 2022 to 2023</u>, November 2023. Note that "literacy" is one component of the EYFS profile, and "communication and language" is another component.



Notes: PBE analysis using EYFS assessment data from the Department for Education. No data collection took place in 2019/20 or 2020/21 due to the Covid pandemic. Note that the EYFS assessment was significantly revised in Sept 2021.

Even if the pre-2021 and current EYFS assessments are not comparable, the evidence of stagnating progress calls for immediate action.⁴ To support children to achieve their full potential, it is necessary to understand the areas most in need of targeted support for children at risk of not meeting expected standards, and the cost implications for those children and wider society.

Areas of deprivation and the need for CLL support

The percentage of five-year-old children not reaching the expected level of CLL development varies significantly across regions in England, as illustrated in Figure 2. In 2022/23, the North West, North East, West Midlands and Yorkshire and the Humber were among the regions with the highest rates of children struggling to meet the literacy standards (32-33%). This amounts to a total of 27,876 five-year-olds achieving below the expected level in literacy in the North West, 8,865 in the North East, 22,117 in the West Midlands and 19,514 in Yorkshire and the Humber.

⁴ The levelling off in the number of children not meeting the expected standards for literacy and communication and language skills could be interpreted as indicating that there is a very small number of children left that could reasonably meet the standard. However, Norbury et al. (2016) suggest that just 10% of children have long term health conditions and disabilities that would impair their ability to reach the expected CLL level. It therefore follows that, between 2018/19 and 2022/23, an estimated 17-20% who could reasonably reach the expected level in literacy did not – the estimate being 8-10% of children for communication and language. These suggest that, far from reaching a natural stopping point, progress has prematurely paused.

The North West, West Midlands, Inner London and North East held the highest rates of children not reaching the expected standard in communication and language (21-23%).

In contrast, the South East recorded slightly lower rates of young children not meeting the expected standard in the same period, with 28% of children for literacy and 17% for communication and language.

Figure 2. Regional inequalities in early years CLL skills persist

Proportion of five-year-olds not reaching the expected level in the literacy component of the EYFS (top) and the communication and language component (bottom) by English region





Deprivation plays a significant role in this geographic spread, because there is a strong link between deprivation and the ability to provide a home learning environment in which children's CLL skills thrive.⁵ Though virtually all parents and care givers want the best for their children and strive to deliver it, lower income families face greater barriers to supporting their children to develop their CLL skills. These include capability barriers, with some parents having less awareness of the importance of communication with very young children and a lack of understanding about the ways to achieve it. Lower income families can experience motivational barriers to supporting their children with literacy, for example, if they lack role models that can teach them the importance of developing their children's literacy skills. They may also be held back by a personal lack of confidence in their own literacy skills. Additionally, lower income families can experience greater opportunity barriers, with limited financial resources restricting the provision of age-appropriate books, toys and other enrichment activities which can aid skills development; cramped conditions acting as a barrier to one-to-one communication; anti-social behaviour in the area making it more difficult to access amenities – of which there may be fewer in the first instance; and lower income parents with multiple jobs or larger numbers of children can also struggle with significant time pressures.⁶

This dynamic is demonstrated further in Figure 3, which shows that the most deprived areas of the country have the highest percentages of children not meeting the expected standards for CLL. While in the least deprived local authorities 28% of five-year-olds did not meet the expected standards for literacy, that percentage rises to 33% in the most deprived local authorities. Similarly, 18% of five-year-olds did not meet the expected standard of communication and language in the least deprived local authorities, but 23% were not meeting the standard in the most deprived local authorities. PBE's analysis found that if CLL skills in every region could be brought up in line with the best performers, then it would more than halve the number of children not meeting the expected level in CLL.⁷

⁵ Hartas, <u>Families' social backgrounds matter: socio-economic factors, home learning and young children's language, literacy and social outcomes, 2011.</u>

⁶ Improving the home learning environment: a behaviour change approach, HM Government and National Literacy Trust, November 2018.

⁷ PBE's analysis classifies Local Authorities (LAs) in three groups based on the average score of the Index of Multiple Deprivation (IMD) in 2019. Therefore, LAs are classified as having "Low" levels of deprivation if they are in the lowest 30% in the distribution of IMD scores, medium deprivation if they are between 30%-60%, and high deprivation above 60%. Ministry of Housing, Communities & Local Government, English Indices of Deprivation 2019, File 11 Local Authority District Summaries (upper-tier), 2019.

Figure 3. Young children are more likely to struggle with CLL in more deprived areas

Proportion of five-year-olds not reaching the expected level in literacy, and communication and language, in local authorities categorised by deprivation





Understanding the growing scale, spread and impact of need

Even though a direct comparison of the total number of children not meeting the expected standard for CLL skills before and after the pandemic is challenging for the reasons outlined above, professionals are confident that the pandemic worsened the state of CLL skills among young children in England. Lockdowns created substantial difficulties; the impact of the closure of many formal early years settings was compounded by the challenges that many parents faced balancing work, financial concerns and supporting their children's learning at home while potentially feeling isolated from their usual support networks.

It is likely that the pandemic exacerbated the gap between the most and least deprived children too. An array of evidence points to educational inequalities worsening during this time. For example, 31% of middle-class parents received online educational support from their provider compared to 23% of working class parents.⁸ Even when early years settings were open, there were large falls in the take-up of funded entitlement places among disadvantaged two-year-olds, with ethnic minorities and children with special educational needs the most likely to have missed out on early years care. In Autumn 2021, early years attendance was still only 90% of what would have been expected without Covid, with attendance lowest in

⁸ The Sutton Trust, <u>COVID-19 and Social Mobility Impact Brief #4: early years</u>, July 2020.

deprived areas and areas with low female employment, suggesting that disadvantaged families and households facing unemployment continued to miss out the most.⁹

With the pandemic and now the cost of living crisis posing great challenges to early years development, effective solutions to address the CLL gaps are paramount.¹⁰ To understand the impact of failing to find those solutions, it is important to begin with the economic costs to children and society if they do not get the help they need.

⁹ La Valle et al, <u>Implications of COVID for Early Childhood Education and Care in England</u>, Centre for Evidence and Implementation, June 2022. There were large falls in take-up of funded entitlement places amongst disadvantaged two-year-olds and, to a lesser extent, amongst three-year-olds, especially in areas with large ethnic minority populations and limited labour market participation.
¹⁰ National Literacy Trust, <u>Children and young people's access to books and educational devices at home during the cost-of-living crisis- A survey of over 3,000 parents and carers in 2023, February 2023.</u>

Our Approach

The economic analysis in this report only focused on the economic costs arising from not sufficiently supporting children to attain the expected level of literacy.

While communication and language are also areas that require focus, a higher percentage of children did not meet the expected standard for the literacy component of EYFS than the communication and language component. This is true when looking at the national, regional or local authority level. Conceptually this makes sense, as evidence suggests communication and language are integral to literacy skills.¹¹

Additionally, it is difficult to isolate the number of children who do not reach the expected level in only literacy, only for communication and language, or both. The Department for Education data on the early years assessment results is presented in a way so that the nature of the overlap of children who do not achieve the expected level in these two areas is unknown. The lower the overlap, the more children are assumed to not be achieving the standard in at least one of these areas.

PBE has chosen to focus only on the conservative scenario in the central analysis results, i.e. it is assumed all children who did not meet the standard for the communication and language component also did not meet the standard for the literacy component. This means that fewer children are assumed to be falling short of the standard in at least one area. This results in a smaller estimate for the number of children who could reasonably meet the standard with sufficient support, and thus provides a conservative estimate for the economic cost of failing to support these children to reach their potential.

Measuring literacy skills in early years

Before exploring the methodology, it is important to note this study uses the EYFS¹² as a measure of children's development in literacy skills. This is the statutory assessment for children's level of development given at the end of the year in which they turn five, usually reception year. This assessment provides a comprehensive summary of the children's level of

¹¹ www.speechandlanguage.org.uk/help-for-families/resource-library-for-families/speech-language-andcommunication-skills-and-literacy/, accessed 12 February 2024.

¹² The early years Stage profile was significantly reviewed in September 2021. Main changes have been made to the wording in the educational programmes, this means including more details on the activities to do with children and a particular focus on examples of activities to develop vocabulary skills across all areas of learning. The areas of learning remain the same; 3 prime areas: communication and language, physical development, personal, social and emotional development and 4 specific areas: literacy, maths, understanding the world and expressive arts and design. Department for Education, Early Years Foundation Stage Profile Results 2022/23, November 2023.

development in 17 Early Learning Goals across seven areas of learning: mathematics, literacy, communication and language, personal, social and emotional development, physical development, expressive arts and design, and understanding of the world. Children are defined as having a Good Level of Development if they are at the expected level in 12 Early Learning Goals within the initial 5 areas of learning. Figure 4 describes how children are assessed in the EYFS.



Figure 4. Early Years Foundation Stage assessment

The economic cost of low literacy skills

This study used existing evidence to establish a link between achieving a Good Level of Development in literacy skills in the early years assessment at age five and academic achievement at Key Stage 1, and the consequent impact on lifetime economic outcomes.

To estimate the economic costs of low literacy skills, PBE took the approach below:

1. Estimate the number of five-year-old children who are not currently achieving the expected standard for literacy but with more tailored support could reasonably achieve the expected level. This analysis estimated the number of children that could be lifted to meet the EYFS standards in literacy. First, data on the number of children not

meeting the standard for the literacy component was taken from EYFS results in 2018/19.¹³ PBE uses evidence from Norbury et al. (2017) to estimate the percentage of five-year-olds who might require special support as they are likely to have persistent Speech, Language and Communication needs (SLCN) and excluded them from the first number, so the remaining group of children were only those that could reasonably achieve the expected level in literacy.

- 2. Estimate the probability of reaching the expected level in Key Stage 1 given the child's performance in CLL in the EYFS assessment. PBE used evidence from Atkinson et al. (2022) to establish the link between having a Good Level of Development in the EYFS assessment and academic achievement in Key Stage 1.¹⁴ Given that this evidence refers to all areas of EYFS and the focus of this analysis is on literacy, PBE's analysis adjusted findings from Atkinson et al. (2022) to isolate the effect of having a Good Level of Development on Key Stage 1 achievement due to literacy alone. Further details about how this analysis used evidence from Atkinson et al. (2022) study is provided in Annex A.
- 3. Use evidence from Paull & Xu (2017) to estimate associations between EYFS performance and later outcomes. Paull & Xu (2017) used data from the National Pupil Database to estimate the link between Key Stage 1 attainment at age 7 and later life outcomes. PBE then linked the Atkinson evidence and Paull & Xu evidence to estimate the probability of later life outcomes given EYFS attainment, via KS1 attainment which is linked to them both.
- 4. Estimate the economic impact arising from these later life outcomes. PBE updated the value of later outcomes in Paull & Xu (2017) to 2022 prices.¹⁵ The value of later outcomes reflects the economic value associated with Special Education Needs recipients, persistent truancy, permanent exclusion, and impact on future probability of employment and wages. As the scope of this analysis includes local authorities, lifetime earnings were adjusted using local wages detailed in the Annual Survey of Hours and Earnings dataset to reflect regional disparities.¹⁶ Further details about how this analysis used evidence from Paull & Xu's (2017) study is provided in Annex B.
- 5. Calculate the total implied economic cost associated with low literacy skills in early years. Multiplying the probabilities in step 3 by

¹³ Early years assessment results. Department for Education, <u>Early Years Foundation Stage Profile</u> <u>Results 2018/19</u>, February 2021.

 ¹⁴ Atkinson et al., 2022., <u>Can holistic school readiness evaluations predict academic achievement and special educational needs status? Evidence from the Early Years Foundation Stage Profile, 2022.
 ¹⁵ GDP deflator last updated, HM Treasury, <u>GDP deflators at market prices and money GDP</u>, 29
 September 2023. Note – 2022 is the latest year with full data available.
</u>

¹⁶ ONS, <u>Earnings and hours worked, place of residence by Local Authority: ASHE Table 8.7a</u>, November 2023.

the economic values in step 4, the cost per child of not reaching the expected level of attainment in EYFS for literacy was estimated. PBE multiplied this cost per child by the total number of children who could reasonably achieve the standard level in literacy defined in step 1 of this methodological approach – this formed the estimate of the overall cost of failing to sufficiently support children to reach their literacy potential.

The analysis in this stage also extended the estimated costs to local authorities in England, using the number of children who did not reach the expected level in literacy by local authority. More details on this are provided in Annex C.

The case of Birmingham

- The number of five-year-old children who are not currently achieving the expected standard for literacy and who, with more tailored support, could reasonably achieve the standard level
 - a. In Birmingham, 15,788 five-year-olds were assessed in the EYFS in 2018/19.
 - b. Of the total number of five-year-olds assessed in the EYFS in 2018/19, 4,820 were below the expected literacy standard.
 - c. PBE calculates that 10% of the total number of five-yearolds (around 1,578 children) will not be able to reach the expected level in literacy as a result of language disorders or other Special Educational Needs and Disabilities. ¹⁷¹⁸
 - d. Subtracting 1,578 children from the 4,820 who did not meet the standard for literacy, PBE's analysis estimates that 3,241 children in Birmingham are not meeting the expected standard in literacy but could reasonably reach it.
- 2) The marginal point improvement in Key Stage 1 given the child's performance in literacy in the EYFS assessment.
 - a. Drawing on previous research by Atkinson (2022), PBE's analysis has estimated that a child in Birmingham who reaches a Good Level of Development in all EYFS areas has a 40% higher chance of performing at expected levels in the Key Stage 1 reading, 32% for writing, 43% for maths and 52% for science. However, it is interesting to know how these probabilities change when only literacy is considered.
 - b. To isolate the effect of a Good Level of Development in literacy alone, PBE estimated that among the total

 ¹⁷ Norbury CF, Vamvakas G, Gooch D, Baird G, Charman T, Simonoff E, Pickles A, <u>Language growth in children with heterogeneous language disorders: a population study</u>, 2017.
 ¹⁸ Department of Health & Social Care and Department for Education, <u>Best start in speech</u>, <u>language and communication</u>: <u>Guidance to support local commissioners and service leads</u>, 2020.

number of children not reaching the expected standard for literacy during the 2018/19 EYFS assessment in Birmingham, only 593 children exclusively failed to meet the standard for literacy. 19 This represents approximately 12% of the total number of children not meeting the standard for literacy in Birmingham.

c. PBE's analysis has adjusted the probabilities of a child in Birmingham reaching the expected level in Key Stage 1 areas due to a Good Level of Development in all areas of the EYFS assessment to isolate the impact of literacy only. This means only taking 12% of the estimated probabilities in point a).

Therefore, a child in Birmingham reaching a Good Level of Development in literacy will get 0.2 additional points spread across all Key Stage 1 areas. PBE has calculated this as follows:

- Reading: 40%* 12%= 5% higher probability of an extra 1 point at KS1. This is equivalent to an average increase in points of 0.05.
- Writing: 32%*12%=4% higher probability for an additional 1 point at KS1. This is equivalent to an average increase in points of 0.04.
- Maths: 43%*12%=4% higher probability for an additional for 1 point at KS1 =0.04. This is equivalent to an average increase in points of 0.04.
- Science: 52% *12%=6% higher probability for an additional for 1 point at KS1=0.06. This is equivalent to an average increase in points of 0.06.

Summing up all Key Stage 1 areas: 0.05+0.04+0.04+0.06 = 0.2 points

- 3) Key Stage 1 performance and later outcomes Evidence from Paull & Xu on the link between KS1 and later outcomes suggests that a one-point increase in the total point score across all Key Stage 1 subjects:
 - reduces the expected duration of SEN support by 0.42 years.
 - reduces the expected length of persistent truancy by 0.009 years.
 - reduces the probability of ever being permanently excluded by 0.3 percentage points.

¹⁹ To calculate this, PBE focuses on the difference between the number of children not meeting the expected standard for literacy (4,820) and maths (4,227) in Birmingham, which are the two areas of the EYFS children were least likely to meet the expected standard.

 increases the probability of achieving 5 good GCSEs (at grades A*-C) by 5 percentage points and increases the probability of achieving at least 1 A level by 7 percentage points.

Considering this evidence by Paull & Xu and PBE's updated estimates of the monetary value of later outcomes associated with a one-point increase across all Key Stage 1 areas, the present value of lifetime outcomes (i.e. when a child is five) for a child in Birmingham in 2022 prices is the following:

- SEN: £1,365
- Truancy: £7
- Exclusion: £97

For lifetime earnings, assuming that a child who studied in Birmingham might work in the local area, PBE's analysis multiplies the average earnings impact associated with a 1 point increase across all KS1 areas by the ratio of the average Birmingham wage to the average UK wage. By factoring in wage differences by region in this way, PBE estimates the impact on lifetime earnings is calculated to be £21,180 per child with a onepoint increase in KS1 overall attainment. Therefore, summing up, the total economic benefits of a child in Birmingham experiencing this marginal increase in KS1 attainment is £22,649.

4) Estimating the lifetime economic cost associated with foregone early years literacy improvement

The 3,241 children in Birmingham not reaching the expected level in early years literacy means they miss out on these economic benefits – in which case, these foregone benefits become costs. Therefore, PBE multiplied the £22,649 lifetime economic costs for a one-point increase across KS1 areas by the estimated 0.2 point increase in KS1 total score associated with achieving the expected level in early years literacy. This estimates that the lifetime economic cost associated with foregone early years literacy improvement is £4,656 per child. This figure represents the cost to society from a child not reaching the expected level in literacy at five years of age, despite having the potential to reach it.

Further multiplying this per child cost by the 3,241 children who PBE estimates are currently not reaching the expected level but could reasonably do so, the analysis suggests that the total cost across an annual cohort of five-year-olds could be as much as £15 million.

Key assumptions

To be able to estimate the implied economic costs of not supporting children to reach the expected level in literacy using the evidence available, PBE considered the following assumptions in the analysis:

- Evidence by Atkinson (2022) defined the Good Level of Development in four additional areas²⁰ apart from literacy. PBE's analysis aimed to find the unique effect of literacy on Key Stage 1 and later outcomes. PBE's analysis used data from the EYFS assessment in 2018/19 - as this is the closest data to the evidence provided by Atkinson et al. (2022) - to calculate the proportion of children that would have passed in all areas if they had passed literacy. Only that proportion is attributed to the estimated change in probability of achieving the expected Key Stage 1 level, to accurately reflect the effect of literacy alone. PBE's analysis tested this assumption by allowing the least overlap possible between EYFS areas. Sensitivity Test 3 explore the impact of this assumption.
- 2. This analysis has assumed that the results from the Atkinson et al. (2022) study, used to link EYFS performance to KSI attainment, are representative of causal relationships for the wider population of today's five-year-olds. There are two key limitations for our purposes: Firstly, it is based on observed data rather than an experiment specifically designed to identify a causal relationship, and secondly, although it uses a large sample of children from the Born in Bradford study, the group is disproportionately more disadvantaged and more ethnically diverse than the wider population of England. However, disadvantage and ethnicity as well as a host of other factors identified as important for academic performance are controlled for in Atkinson et al.'s analysis. This gives us confidence that the marginal effects used in the analysis in this paper should broadly represent causal relationships for the wider population. The impact of alternative assumptions for this link is explored in Sensitivity Test 1.

²⁰ These areas are communication and language, physical development, personal social and emotional development, and maths.

- 3. This analysis also assumed that at least 10% of the total number of five-year-olds would need special support to increase their literacy skills as they are likely to have SLCN²¹. An alternative scenario to test this assumption was also estimated and can be found in the Sensitivity Test 2.
- 4. This study assumed that the relationship between the old EYFS and KS1 performance is broadly representative of the likely relationship between the new EYFS and KS1 performance. PBE's analysis reduced potential concerns from this issue by calculating the total implied costs using results from the previous EYFS (i.e. 2018/19).
- 5. The analysis assumed that the relationship between KS1 performance and future outcomes established for the average child in Paull and Xu (2017) is representative of the likely outcomes for children that did not reach the expected level in the literacy component of the EYFS in 2018/19.
- 6. When calculating lifetime earnings at the local area, PBE's analysis factored in differences in wages for each local authority. The assumption behind this was that children studying in one area as five-year-olds will be employed in that area as adults. This is a simplified assumption as some individuals will move to work in different areas. This will impact benefit estimates at the local authority level but not at the aggregated national level.
- 7. This analysis has not incorporated any "general equilibrium effects". In a theoretical example where all individuals suddenly received support to improve their CLL skills, this may plausibly lead to a reduction in the wage premium for these skills (and effectively reduce the scale of the economic benefits identified in this paper). The scale of this effect is uncertain – in a small open economy such as the UK the effects could be relatively limited, provided that the economy re-structures over time towards high-skill, high-wage sectors. The analysis in this report is focused on the benefits to cohorts of children that will initially make up a small proportion of the total labour force – it would take 40-50 years before the effects of increased early-years CLL skills filter through to the entire labour market. This provides time for the economy to adapt. A similar process was demonstrated in an Institute of Fiscal Studies analysis of historical changes in skill levels in the UK which highlighted that the proportion of people with a university degree by age 30 more than doubled between those born in 1965-69 compared to those born 10

²¹ Norbury CF, Vamvakas G, Gooch D, Baird G, Charman T, Simonoff E, Pickles A, <u>Language growth in</u> <u>children with heterogeneous language disorders: a population study, 2017</u>. Department of Health & Social Care and Department for Education, <u>Best start in Speech, language and communication</u>: <u>Guidance to support local commissioners and service leads</u>, 2020.

years later. However, the high skill wage premium remained largely unchanged in the UK over this period as firms adapted their production approach to make use of the more highly skilled labour force.²² As such, it is unlikely that wider "general equilibrium effects" will have a substantial impact on the results reported in this study.

8. Even though models estimated by Atkinson et al. (2022) on the relationship between EYFS attainment and Key Stage 1 outcomes account for several factors on children's background such gender, ethnicity, academic month of birth, pre-term status, maternal education, whether a child had English as an additional language and free school meals status, other factors were not considered in their models due to lack of data and could have potentially affected the results.²³ For example, the fact that results from EYFS and Key Stage 1 rely on teacher-reported performance could directly introduce a bias, it could be the case that teachers have an implicit bias for children of certain ethnic-minority backgrounds which in turn affects the assessment of their performance.²⁴

Additionally, any benefits of children not attaining the expected level of literacy in the EYFS assessment are not included in the scope of this study. While successful interventions may incur considerable costs, e.g. to cover the staff, materials and indirect costs of delivering a tailored literacy support programme, these are not currently defined and quantified. Therefore, this study does not claim to provide a balanced cost-benefit analysis of a given proposed intervention to improve early years literacy levels. Instead, it serves to explore the potential scale and cost of a problem to understand the need for a solution.

²² Blundell R, Green D, Jin W, <u>The UK wage premium puzzle: how did a large increase in university</u> <u>graduates leave the education premium unchanged?</u> Institute for Fiscal Studies Working Paper W16/01, 2016.

²³ Atkinson et al 2022 also highlighted this as a limitation of their study. The authors also recognise that when estimating links between EYFS assessment and Key Stage 1 outcomes, teacher demographics are a factor that should be controlled for.

²⁴ This is based on this previous research by Gilliam et al 2016: Yale Child Study Center, <u>Do Early</u> Educators' Implicit Biases Regarding Sex and Race Relate to Behaviour Expectations and Recommendations of Preschool Expulsions and Suspensions? September 2016.

Results of the analysis

National findings

The economic cost of insufficient support for early years literacy is likely to be significant. The lifetime economic cost for the typical child that does not meet the expected standard for literacy at age five, but could with adequate support, is estimated to be £7,800. This results from the knockon effects that lower literacy skills at a young age are estimated to have on subsequent school engagement, academic success and later employment outcomes.²⁵

These costs accrue to both the individual and to government. £5,300 (68%) of the lifetime costs are incurred by the individuals themselves due to lower earnings resulting from the reduced probability of being employed and lower pay whilst in work. The remaining £2,500 (32%) of the lifetime costs fall to the government, as a result of increased education, social and welfare spending and reduced tax receipts.

While this cost is notable for each child involved, the costs at a countrywide level are even greater. Our analysis suggests there are around 106,000 five-year-old children each year who aren't currently meeting the expected standard but could with adequate support.²⁶ The total lifetime costs of not meeting the EYFS literacy standard for this entire group are £830 million. Further details on PBE's methodology for these calculations can be found in Annex A, B and C while Figure 5 illustrates this calculation.

²⁵ Based on PBE's updated values assigned to the economic impact associated with a marginal effect in KS1 attainment. In Annex B, Table 5, this is illustrated in the costs arising from a one-point reduction in KS1 attainment totalling £25,172 per child. Of this, £17,023 (or 68% of £25,172) falls to the individual, and £8,149 (32% of the total) to the government.

²⁶ This analysis is based on the assumption that 10% of the 639,000 children assessed at the EYFS in 2018/19 would struggle to meet the EYFS standard without highly specialised support due to persistent Speech Language and Communication Needs. These 63,900 children are excluded from the total of 172,000 children that did not meet the EYFS standard for literacy in that year, leaving 106,000 children that did not meet the standards but realistically could have with some additional support.

Figure 5. Estimated cost for England



£269 million of this cost is incurred by the government. This includes £49 million in additional education and welfare services related to SEN support, truancy and exclusion and £220 million in fewer taxes and national insurance contributions. The remaining £561 million represents the impact on individuals in the form of reduced lifetime earnings.

Regional and local authority estimates

The costs of not meeting the expected literacy standard for five -year-olds are unlikely to be spread evenly across different areas of England. The South East, North West and East of England regions accumulate the highest implied economic costs of low literacy as shown in Figure 6 – with low literacy costing the South East region around £150 million for each cohort of five year olds, the North West approximately £110 million and the East of England approximately £100 million.

Figure 6. Failing to support children's literacy costs generates the largest cost for South East

Total implied economic costs, by region, of each annual cohort of five-year-olds not reaching the expected level in literacy even though they reasonably could



Notes: PBE analysis using various sources including Department for Education (2023), Atkinson et al (2022), Paull and Xu (2017).

Regional differences in the total estimated long-term costs of children not meeting the EYFS literacy standard are driven by a range of factors including: the total size of the population in an area, the share of children below the expected level in literacy, and average local earnings. The South East, for example, has a large population and relatively high wages and therefore experiences large economic costs from poor literacy, though it is not one of the worst performing regions.

However, given the strong links between literacy skills and deprivation it is valuable to focus on the top ten most deprived local authorities in England where the need for additional support is likely to be greatest. The results highlight the significant economic costs of not supporting children to reach the standards in literacy in Birmingham, Liverpool, and Manchester. In the top 10 most disadvantaged local authorities in England, approximately 3 out of every 10 five-year-olds are not currently meeting the expected standard in literacy. Not supporting these children to meet the expected standard could lead to overall economic costs of £47 million per school year, as shown in Table 1.

Deprivation rank	Local authority	Region	% of children not meeting the EYFS standard for literacy	Total implied economic costs (Millions)
1	Blackpool	North West	30%	£1.44
2	Knowsley	North West	31%	£2.44
3	Liverpool	North West	34%	£7.63
4	Kingston Upon Hull City of	Yorkshire and the Humber	32%	£3.76
5	Middlesbrough	North East	36%	£2.06
6	Manchester	North West	33%	£7.10
7	Birmingham	West Midlands	31%	£15.09
8	Blackburn with Darwen	North West	30%	£1.84
9	Hartlepool	North East	26%	£0.76
10	Nottingham	East Midlands	33%	£4.77

Table 1. Top ten local authorities by deprivation

Notes: PBE analysis using various sources including Department for Education (2023), Atkinson et al (2022), Paull and Xu (2017).

Full details of the total implied economic costs for other Local Authorities can be found in the spreadsheet Annex D.²⁷

Sensitivity analysis

A range of alternative scenarios were examined to test the robustness of these results. They highlight that the broad conclusions of the analysis - that there is a substantial long-term economic cost from failing to support children to meet the expected standard for literacy at age five – are robust to variations in the key assumptions. The estimated total cost for a single cohort of children is likely to be in the range of £0.5 billion to £3.5 billion.

The following scenarios are explored:

- Sensitivity Test 1: Alternative scenarios for the link between early years assessment and Key Stage 1 performance.
- Sensitivity Test 2: Reduced proportion of children with persistent Speech Language and Communication Needs.
- Sensitivity Test 3: Minimum overlap possible between those not meeting the expected standard for different subject areas of the EYFS.
- Sensitivity Test 4: Using alternative assumptions for the link between KSI attainment and lifetime economic benefits.

²⁷ Annex D is available at https://www.probonoeconomics.com/early-literacy-matters.

Sensitivity Test 1: Alternative scenarios for the link between early years assessments and Key Stage 1 performance.

Alternative assumptions are used for the estimated odds of reaching the expected level in KSI, given a Good Level of Development in EYFS., based on the levels of uncertainty provided by Atkinson et al. in their analysis – known as the 95% Confidence Interval. Further details of Atkinson et al. (2022) can be found in Annex A.

When taking the lower bound of the odds, the cost per child over their lifetime increases to £ 8,989 which translates to total costs of £953 million. In comparison, when taking the upper bound of the odds the cost per child over their lifetime falls to £4,747 cost per child over lifetime. This translates to total costs of £503 million per school year group.²⁸

Sensitivity Test 2: Reduced proportion of children with persistent Speech Language and Communication Needs.

In the core results, it was estimated that 10% of the total number of fiveyear-old children (63,899 children) might need significant specialist support to be able to reach the expected level of literacy due to persistent Speech Language and Communication Needs based on Norbury et al. 2017.²⁹ These children were excluded from the analysis.

In the sensitivity analysis, it is assumed that the percentage of children who are not able to reach the expected level of literacy or need special support to do so might potentially be as low as $7\%^{30}$ (a total of 44,730 children). In this scenario, the estimated total costs increase to approximately £987 million.

Sensitivity Test 3: Minimum overlap possible between those not meeting the expected standard for different subject areas of the EYFS

The core analysis adjusted the estimated difference in probabilities of reaching the expected level in Key Stage 1 by estimating the number of children who would have passed all areas of EYFS if they had passed the literacy standard. This calculation assumed a maximum overlap between the two prime EYFS areas where children were most likely to not meet the

²⁸ Note that the upper and lower bound estimates in this sensitivity test are not symmetrical as they reflect the asymmetrical Confidence Intervals of the odds ratio estimates in Atkinson et al 2022.By definition, confidence intervals of Odds ratio (OR) do not have symmetry because of the fact the OR is skewed to the right (being ranged between 0 and ∞). The asymmetrical nature of the odds ratio has been discussed in : Bland and Altman, <u>The odds ratio</u>, 2000 and Islam, <u>Symmetry of Odds Ratio</u>, 2013. ²⁹ Norbury CF, Vamvakas G, Gooch D, Baird G, Charman T, Simonoff E, Pickles A (2017): <u>Language growth in children with heterogeneous language disorders: a population study</u>. Journal of Child Psychology and Psychiatry, 58(10), 1092-1105.

³⁰ Based on House of Commons Library, <u>Speech, language, and communication support for children</u>, 2018.

expected standard (i.e. literacy and maths). Full details of this calculation are available in Annex A.

In this sensitivity analysis, the core results are tested by assuming the least overlap possible between literacy and maths. This increases the strength of the relationship between meeting the EYFS literacy standard and KS1 by around 40%.³¹ This results in an average implied economic cost of £10,995 per child and a total estimated cost of £1.2 billion per cohort.

Sensitivity Test 4: Using alternative assumptions for the link between KS1 attainment and lifetime economic benefits

The analysis in this report links the relationship between EYFS assessment and Key stage 1 outcomes from Atkinson's (2022) to the associations between Key Stage 1 and later outcomes in Paull & Xu's (2017). Our core scenario used Paull & Xu's main measurement of KS1 attainment - "Total points in all Key Stage 1 subjects".^{32 33} However, Paull & Xu provide several alternative sets of results that use different definitions of KS1 attainment.

This sensitivity analysis replicates the core results with three alternative measures of KS1 attainment provided in Paull & Xu: "achieved expected level in main KS1 subjects", achieved expected level in all KS1 subjects" and "Total points in main KS1 subjects", and is presented in Table 2. These alternative scenarios highlight that total costs could increase substantially.

Measures	Average implied economic cost per child	Total implied economic costs (Millions)
Achieved expected level in main KS1 subjects	£32,803	£3,479
Achieved expected level in all KS1 subjects	£31,758	£3,368
Total points in main KSI subjects	£12,012	£1,274
Total points across all KSI subjects (PBE's core analysis)	£7,822	£830

Table 2. KSI attainment measures in Paull & Xu (2017) and PBE results

³¹ In this scenario it is assumed that the link between meeting the EYFS literacy standard and KS1 attainment is 27% as strong as the relationship identified for the link between meeting the standard across all EYFS areas and KS1 attainment in Atkinson et al (2022). This compares to 19% in the core scenario. This is based on assuming that an additional 13,054 children did not meet the standard across all areas due to their performance in maths rather than literacy.

³² Paull & Xu, consider five KSI subject areas: speaking and listening, reading, writing, mathematics, and science. Reading, writing, and mathematics are defined as "main subjects". In contrast, Atkinson's study exclusively reports results for four KSI areas: reading, writing, mathematics, and science.

³³ Apart from the predictive power, Paull & Xu's chose the "total point score across all subjects" as the main measure because it allowed for substantial variation between pupils in the NPD sample. The authors also stated that results from their regression models on the associations between KS1 and later outcomes yielded in qualitatively similar results for all four measures of KS1 attainment.

Implications of the findings

The significant long-term economic costs when children do not meet the expected standard for early years literacy highlight the need for further action. However, the complex interactions between deprivation, the home learning environment and literacy mean that tackling this issue will not be easy. There will be a need for targeted interventions focused on those areas of the country most impacted, in a way that builds on the support provided in formal educational settings by reaching into the community.

This is an approach that the National Literacy Trust is taking. They are working in two-fifths of the most deprived Local Authorities in the country, as indicated in Figure 7, targeting areas with some of the lowest proportions of five-year-olds meeting the expected standard for literacy. They build localised networks to bring together early years settings, schools, community groups, businesses and government to ensure children have the communication, language and literacy skills they need to thrive, supporting children from birth. This includes work in Blackpool, the most deprived area of the country where implied economic costs of low literacy are in the region of £1.4 million per cohort of five-year-olds. It also includes Middlesborough, the fifth most deprived area of the country with one of the highest rates of low literacy among young children – with 36% of children not meeting the expected standard for literacy.

Jayden's story

Like many children born just before the Covid-19 pandemic, Jayden was still struggling to speak at the age of three. He was growing frustrated and had started to become violent towards his mum, who didn't know where to turn for help.

Jayden and his mum took part in a National Literacy Trust language skills programme, which completely transformed his development. They now enjoy reading together and Jayden has started speaking, while his behaviour has improved dramatically too. Jayden's mum has noticed a huge change:

"Being able to sit with him and enjoy that time while having a little snuggle...compared to how he was before, it's just amazing." Figure 7. The National Literacy Trust is focusing on those areas that have a high proportion of children not meeting the age-five literacy standard



Notes: Local Authorities Upper Tier (2019). Map displays the % of children not meeting the literacy standard in 2018/19. Orange points indicate Local Authorities where National Literacy Trust works.

If the National Literacy Trust's approach is successful, then the benefits could be enormous. The current cost of failing to support children to meet the literacy standards at age five in the local authorities that the National Literacy Trust is focused on is more than £100m for each cohort of children they work with.

Importantly, this work could also help future generations too. There is strong evidence that educational outcomes play an important role in the "intergenerational transmission of disadvantage" – the process by which the life outcomes of people born into disadvantage are worse, which impacts their children's life outcomes too. Breaking this cycle now could have impacts that reverberate through communities for generations to come, ultimately playing a part in supporting communities to thrive.

Conclusion

England continues to have too many children who struggle to reach the age-five expected standards for communication, language and literacy skills, and this has substantial economic implications. This report estimates that failing to sufficiently support a typical cohort of five-year-olds to meet the expected standard for literacy could generate lifetime economic costs of around £830 million.

With both the pandemic and the cost of living crisis likely worsening the scale of poor literacy among young children, new approaches and strategies to tackle this challenge are essential. But change is possible. The number of children not meeting the expected standard literacy failure rate could be more than halved if the worst performing areas of the country were supported to reach the standards of the best performing areas. Achieving this change requires a better understanding of the role that deprivation and the home learning environment play in some of the worst performing areas. Designing strategies that reach beyond the early years system and into the community will be critical.

No single early years service can do this alone – it will require a community partnership that combines the best of the public, private and social sectors to make a difference. Combining the expertise and structure of the early years sector with the reach into the community provided through social sector organisations and, importantly, employers could offer a new way to unleash the potential of the next generation of learners.

The 'Early Words Matter' campaign led by the National Literacy Trust provides an excellent example of what this new approach could look like.

Annex A –The relationship between EYFS and Key Stage 1- details on Atkinson (2022)

This section provides further details of the evidence from Atkinson et al. 2022, which was drawn to establish a relationship between the Early Years Foundation Stage (EYFS) assessment and later academic achievement in Key Stage 1. This evidence provides insights on how the Good Level of Development measure³⁴ derived from the EYFS assessment predicts academic achievement at the end of Key Stage 1 using data from children in the Born in Bradford longitudinal sample.³⁵

We approach this evidence following the process below:

• **Part 1:** Turns odds ratio into probabilities from the regression models predicting changes in Key Stage 1 attainment based on meeting the expected standard for EYFS assessment.

A set of models controlling for background and school characteristics³⁶ were estimated in Atkinson et al. 2022 to examine whether reaching the Good Level of Development predicted performance on the Key Stage 1 outcome in four subject areas: reading, writing, maths, and science. Results from these models were expressed in the form of odds ratios.³⁷

To link changes in Key Stage I due to improvements in the EYSP assessment from Atkinson 2022 to estimates on the associations between KSI and the value of later outcomes in Paull & Xu's (2017), odds ratio obtained in Atkinson et al., 2022 needed to be converted into probabilities, which was done following two steps:

A) Obtaining coefficients by calculating the natural logarithm of the odds ratio reported for each Key stage 1 outcome as: ³⁸

$$\beta_{x=1,2,\dots,K} = \ln(odds_{x=1,2,\dots,K})$$

Where β is the estimated coefficient for variable x and $\ln(odds_x)$ is the natural logarithm of the odds ratio for variable x.

³⁴ The Good level of Development measure was defined as a binary variable: "reached" or "not reached". ³⁵ The Born in Bradford longitudinal study contains information on 13,858 children born between 2007 and 2011. However, the analysis on Key Stage 1 academic achievement was restricted to a sample of 5,777 children, as outcomes on Key Stage 1 were fully completed for these children.

³⁶ Models account for background characteristics such as gender, SEN status, academic month of birth, maternal education, ethnicity and whether the child received free school meals or spoke English as an additional language during reception year.

³⁷ For each Key Stage I subject area, the authors calculated two set of models: "below expected vs expected" and "expected vs above expected", this analysis uses odds-ratio from the former.

³⁸ Presents odds-ratios in figures by Key Stage 1 outcome, results in this analysis refer to Figure 4 and Figures D1, E1, F1 in the supplementary material of Atkinson et al., 2022 respectively. <u>Can holistic school</u> readiness evaluations predict academic achievement and special educational needs status? Evidence from the Early Years Foundation Stage Profile, 2022.

B) Once coefficients were obtained, they were turn into probabilities using the following formula³⁹.

$$\Pr(Y=1 \mid X1, X2, \dots, X_k) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X1 + \beta_2 X2 + \dots + \beta_k X_k)}}$$

Where P_r is the estimated probability of reaching the expected level in Key Stage 1 for a given subject (i.e. reading, writing, maths or science). *Y* is the outcome of the EYFS assessment which takes the value of 1 for "Good level of development reached "or 0 for "Good level of development not reached yet". *X* is the average value⁴⁰ for a given variable and $\beta_{x=1,2...K}$ are the coefficients obtained in step A.

In terms of the β_0 , which represents the constant in the regression models, PBE had to assume that the constant was zero for all models as it couldn't be calculated as the other regression coefficients.

• Part 2: Estimate changes in predicted probabilities of performing at expected level in Key Stage 1 by subject.

Difference in predicted probabilities were calculated for each Key Stage 1 subject as follows:

$$Diff_s = P_r | Y = 1 - P_r | Y = 0$$

Where the *Diffs* is the change in probabilities for subject $s=\{1,...,4\}$ (i.e. reading, writing, maths or science) from "Good level of development reached " to "Good level of development not reached yet". $P_r|Y = 1$ is the probability of reaching the expected level in Key Stage 1 in $s=\{1,...,4\}$ given that the "Good level of development has been reached". $P_r|Y = 0$ is the probability of reaching the expected level in Key Stage 1 in $s=\{1,...,4\}$ given that the "Good level of development has been reached". $P_r|Y = 0$ is the probability of reaching the expected level in Key Stage 1 in $s=\{1,...,4\}$ given that the "Good level of development has not been reached yet".

Overall, PBE's analysis has estimated that having a Good Level of Development in all EYFS areas results in 1.67 points across all Key Stage 1 areas. The next step looks at how PBE's analysis adjusted this to reflect the impact of literacy alone.

• Part 3: Adjusts the impact to literacy only.

Considering that the change in probability calculated in part 2 refers to all areas of EYFS, this part of the analysis adjusts the estimated change

³⁹ The predicted probability is estimated using the formulae proposed by Stock and Watson -Introduction to Econometrics. (Page 443).

⁴⁰ In this analysis we have taken average values from the Born in Bradford Sample as reported in Atkinson et al. 2022 (Table 1, page 48).

in probabilities to accurately reflect the effect of literacy alone. To make the adjustment, the analysis is reduced to two areas where more children did not meet the expected standard in the EYFS (i.e. literacy and maths). PBE calculated how many children would have passed in all areas of EYFS if they had all passed literacy. This was done following four steps:

- a) Taking the number of children below the expected level in literacy in EYFS 2018/19. Taking the total number for England: 169,958.
- b) Taking the number of children below the expected level in Maths in EYFS 2018/2019: 137,697 total number for England.
- c) Calculating the number of children who did not meet the expected standard for literacy alone as the difference between (a) and (b): 32,261.
- d) Calculating the share of children who did not meet the expected standard for literacy alone over the total number of children who were below the expected level in literacy. This is (c) over (a), which is 19%.
- e) Calculating the effect of literacy on all Key Stage 1 areas: This involves taking the total points impact at KS1 for passing all EYFS (1.67) and multiply that by (d) 19%= 0.32

Annex B – Details on SEED value for Money Paull & Xu (2017)

This Annex provides details of the Study for Early Education and Development (SEED) by Paull & Xu (2017)⁴¹ which is used in this analysis to establish a link between Key Stage 1 and later outcomes. We review evidence by Paull & Xu's in two parts:

Part 1: Link between Key Stage 1 at age 7 and later outcomes in a child's life

Paull & Xu's work uses data from the National Pupil Database from three cohorts of seven-year-old pupils who completed Key Stage 1 between 2002 and 2004. Results are taken from five Key Stage 1 subjects: speaking and listening, reading, writing, mathematics, and science. Each Key Stage 1 subject is assessed in a scale of 0 to 4, where 2 is the expected level of attainment.

Paull & Xu estimate the link between KS1 attainment and later life outcomes using the total point score across all Key Stage 1 subjects⁴². To link Key Stage 1 attainment at age 7 and later outcomes, Paull and Xu estimate a set of regression models that account for children's background characteristics including ethnicity, gender, SEN status prior to KS1 results and socioeconomic status. These regression models estimate the relationship between a one-point increase in the total point score across all KS1 subjects and the following later outcomes:

- The expected duration of SEN support
- Number of years with persistent truancy
- The probability of ever being permanently excluded from school
- The probability of achieving 5 good GCSEs (at grades A*-C)
- The probability of achieving at least 1 A-level

Table 3 presents the change in the probability of each later outcome that is associated with a one-point increase in the total point score across all KS1

⁴¹ Paull and Xu, Study of Early Education and Development (SEED): <u>The potential value for money of</u> <u>early education</u>, 2017.

⁴² Paull and Xu (2017) also present summary statistics for other measures of attainment in Key Stage 1 (KS1) such as: "achieved Level 2 in main KS1 subjects", "achieved level 2 in all KS1 subjects", "total points in main KS1 subjects" and "total points in all KS1 subjects". However, authors preferred the later as they claimed it was the measure that allowed for substantial variation between pupils, and it has also the highest predictive power in terms of the level of qualification achieved and consequently in later outcomes such earnings.

subjects. These figures are based on the results from regression models presented in Table 10 in Paull and Xu's (2017) study.⁴³

Table 3. Summary of estimated impacts of improvements in Key Stage 1 attainment on later outcomes

Later Outcomes	Marginal effect for total points in all KSI subjects found in Paull and Xu's Table 10
Number of years SEN	-0.421
Number of years serv	0.121
Number of years with persistent truancy	-0.009
Proportion of pupils with a permanent exclusion	-0.003
Highest qualification: fewer than 5 GCSEs at A*-C	-0.05
Highest qualification: at least 1 A level	0.072

Even though the total point score across all KSI subjects measure used by Paull & Xu in their main results is not equivalent⁴⁴ to the "expected level in KSI reached" from Atkinson 2022- which would be ideal to make a link between those two pieces of evidence- we have considered the total point score across all KSI subjects measure from Paull & Xu in our core results for four reasons below:

- It is the preferred measure by Paull & Xu, as it allows for substantial variation between pupils observed in the National Pupil Database sample used to drawn associations in their study.
- It has the highest predictive power in terms of the level of qualification achieved and therefore earnings in Paull and Xu's study

 earnings are the largest component of the estimated economic impact of reaching the expected level in early years literacy.
- The "achieved expected level in main KSI subjects" and "Total points in main KSI subjects" measures from Paull & Xu only refer to reading, writing and maths as "main subjects". For this analysis, having all areas is important as previous research have demonstrated the spillover effects from improvements in literacy in overall academic performance. ⁴⁵ Therefore, a definition that considers all areas is

⁴³ PBE's analysis focuses on results from the most comprehensive model (i.e. model that controls for background characteristics of the child which are reported in column 2 of Table 10 in Paull & Xu's (2017)).
⁴⁴ This is because of the areas of KS1. The "achieved expected level in all Key Stage 1 subjects" and "total point score across all KS1 subjects" measures from Paull & Xu's work considers five areas of KS1, which are: speaking and listening, reading, writing, mathematics, and science. Whereas results from Atkinson et al 2022 are only given for four subject areas: reading, writing, mathematics, and science.
⁴⁵ Cimmiyotti, 2013, Impact of Reading Ability on Academic Performance at the Primary Level Primary Level.

more likely to reflect the effects from early literacy skills on KS1 attainment and further in later outcomes.

• Paull & Xu stated that the relationship between KSI attainment and later outcomes yielded qualitatively similar results for all four measures of Key Stage 1 attainment.⁴⁶

PBE's analysis includes a sensitivity test using results of the associations between KSI attainment and later outcomes for the three remaining measures of KSI attainment provided in Paull & Xu (2017) which can be found in the results section.

Part 2: Update the monetary value of improvements in later outcomes associated with improvements in Key Stage 1

To calculate the monetary value of outcomes, Paull & Xu (2017) used data from multiple sources. The monetary value of reductions in the prevalence of SEN and truancy associated with Key Stage 1 attainment are calculated using valuations from the Department for Education and Brookes et al respectively. For the value of SEN, Paul & Xu (2017) estimate the annual cost per child using data from the Department for Education on the annual SEN spending in 2015 and the total number of pupils with SEN on the school roll in the same year. The monetary value of truancy includes the cost of education and welfare services only and is derived from evidence by Brookes et al (2007).⁴⁷

To associate monetary values to the probability of ever being excluded, Paul & Xu (2017) rely on previous research by Brookes et al (2007), which includes the costs to the education system discounted to the age of six and a greater use of social services up to age 16. Other related administrative costs assumed by the Local Authorities are also included.

Finally, valuations of lifetime earnings in Paull & Xu (2017) rely on previous research by Cattan et al (2014), which is considered by the authors to be the most robust evidence for returns to qualifications as they use micro data from the British Household Panel and Labour Force Survey to estimate the impact of educational attainment on expected lifetime gross earnings. Given that estimations from Cattan et al (2014) are presented in 2013 prices

36

⁴⁶ Paul and Xu also estimated additional models for KS1 attainment and later outcomes for three different measures of KS1 attainment: Achieved expected level in main KS1 subjects; achieved expected level in all KS1 subjects and total points in main KS1 subjects. They present summary statistics for all measures in Table 8 in the SEED value for Money report.

⁴⁷ Even though other related costs such lost earnings, health, crime, and other services are reported in the original source by Brookes et al 2007. Paull and Xu (2017) decide to exclude those as they claimed, these related costs are more directly captured in other sources.

and discounted at age four, Paull & Xu (2017) uprate these estimates to 2015 prices and discount values to age three, which is the reference age the authors consider in their early education study.⁴⁸

PBE's approach to monetary values was first looking for any credible sources, starting with the original ones reported in Paull & Xu's study. When original sources were not available, PBE updated the monetary values from Paull & Xu's (2017) study. As the focus of PBE's analysis is to understand the effect of early years education measured through the EYFS assessment at age five on later outcomes, all relevant monetary values are updated to 2022 prices and discounted to age five using a discount rate of 3.5%. Updating later outcomes is explained in detail below:

- SEN: the Department for Education published its total annual SEN spending for 2022/23. PBE directly estimates the average cost of SEN as the ratio between the total costs for SEN in 2022/23 and the total number of children with SEN in 2022/23. This average annual cost is further discounted to the age of five.⁴⁹ Costs for SEN include costs in SEN schools and SEN services at other schools including funding for SEN within an individual school's budget, direct payments for SEN, additional high needs funding for mainstream schools and academies and transport for children with SEN.
- **Truancy:** Brookes et al (2007) estimates the annual cost associated to education and welfare services due to truancy from ages 11 to 16. PBE has updated the lifetime value reported in Brookes et al (2007) to 2022 prices and discounted to the age of five.
- Exclusion: Brookes et al (2007) estimates the costs of exclusion, including the costs of alternative education and social services from the age 13 to 16 and administration costs for the local authority every time a new exclusion case emerge, which according to Brookes is normally at the age of 12. This analysis uprates these costs to 2022 prices and discount them to the age of five.
- Lifetime earnings: for the value of lifetime earnings, PBE updates the monetary valuations reported in Paull & Xu's (2017), which include the value of individual's net earnings and estimated income tax and national insurance revenues for the government associated with five good GCSE's and at least one A level. This analysis takes

⁴⁸ In the Value for Money approach employed by Paull & Xu, the monetised value of all benefits was discounted to the period of impact, specifically when the child reached the age of three or four. This adjustment was made to account for the principle that the value of £1 in the future is lower than its present value. A discount rate of 3.5 percent was implemented, aligning with the related guidelines by HM Treasury (HM Treasury, 2003).

⁴⁹ The total annual SEN spending is taken from section 251 returns (table 2, DfE (2022b)) and the number of pupils with SEN on the school roll (table 1C, DfE (2022)).

those values and uprate them to 2022 prices and present them in present value for a five-year-old.

Details on the estimated values for each later outcome in Paull & Xu's study and the updated values by PBE's analysis are presented in Table 4.

Table 4. Summary of reported monetary outcomes in Paull & Xu and PBE's analysis

Later Outcomes	Paull & Xu (2017) SEED Value for Money Study		PBE's analysis
	Detail period/ discounted age in Table 12 Paull and Xul	Monetary values of outcomes in 2015 prices in Table 12 Paull and Xu	Monetary values of outcomes in 2022 prices discounted at age 5
		(2)	(3)
services) (govt)	Annual Cost	£4,190	±3,243
Truancy: annual cost (education+welfare services) (govt)	Annual Cost	£875	£766
Exclusion: cost per case (LA admin, alternative education + social services (govt)	Cost per case/ discounted to age 3	£27,828	£32,226
Lifetime earnings: 5 good GCSEs- Net earnings (private)	Discounted to age 3	£75,532	£96,527
Lifetime earnings: 5 good GCSEs - Income tax and NI revenue (govt)	Discounted to age 3	£29,639	£37,877
Lifetime earnings: at least one A level - Net earnings (private)	Discounted to age 3	£132,555	£169,399
Lifetime earnings: at least one A level - Income tax and NI revenue (govt)	Discounted to age 3	£52,015	£66,473

Table 5 presents the associated value of later outcomes to Key Stage 1. PBE's analysis calculated this by multiplying the updated monetary value of outcomes presented in Table 4, column 3 by the marginal effect for total points in all KSI subjects found in Paull & Xu's study and reported in Table 3 above.

Later Outcomes	Paull & Xu (2017)	D17) PBE's analys	
	Marginal effect for total points in all KSI subjects	Monetary values of outcomes in 2022 prices discounted at age 5	Associated monetary value 2022 prices discounted at age 5
	(1)	(2)	(1*2)
SEN: annual cost (school+other services) (govt)	-0.421	£3,243	£1,365
Truancy: annual cost (education+welfare services) (govt)	-0.009	£766	£7
Exclusion: cost per case (LA admin, alternative education + social services (govt)	-0.003	£32,226	£97
Lifetime earnings: 5 good GCSEs- Net earnings (private)	0.05	£96,527	£4,826
Lifetime earnings: 5 good GCSEs - Income tax and NI revenue (govt)	0.05	£37,877	£1,894
Lifetime earnings: at least one A level - Net earnings (private)	0.072	£169,399	£12,197
Lifetime earnings: at least one A level - Income tax and NI revenue (govt)	0.072	£66,473	£4,786
Total govt			£8,149
Total private			£17,023
Grand total			£25,172

Table 5. Associated Monetary value of outcomes associated to a one-point increase in total Key Stage 1 attainment.

Annex C – Details of cost calculations at the local authority level

This annex provides an overview of the process PBE has taken to estimate costs at the local authority level.

1. Taking the number of children below the expected level in literacy by local authority.

Data from the Department for Education on the number of pupils below the expected level in literacy in the academic year 2018/19 was extracted by local authority.

2. Calculating the number of children that could reasonably reach the expected level in literacy with more tailored support.

To estimate the number of children that could reasonably reach the expected level, this analysis has estimated the number of children that may have severe language disorders and long term health conditions which would make it difficult or impossible to achieve the expected level without specialist, intense support.

This calculation was done following the process below:

- Taking the number of five-year-olds assessed in the Early Years Foundation Stage (EYFS) by local authority.
- Previous research by PBE has estimated that, according to evidence from Norbury (2016)⁵⁰ and Paull and Xu (2017)⁵¹, 10% of all five-year-olds have significantly low vocabulary assessment skills or did not have the capacity to take the assessment due to their severe health or disability.52PBE's analysis has estimated this proportion of children by local authority.
- These children have, by definition, such low or no attainment (due to severe SEND and long-term health conditions) in early years assessment that indicates they need specialist support

⁵⁰ Norbury et al (2016), <u>The impact of nonverbal ability on prevalence and clinical presentation of</u> <u>language disorder: evidence from a population study.</u>

⁵¹ Paull and Xu, Study of Early Education and Development (SEED): <u>The potential value for money of</u> <u>early education</u>, 2017.

⁵² PBE, <u>The economic cost of early vulnerable language skills, July 2021</u>. According to this paper, the 10% estimate consists of 7% of three-year-olds in the SEED dataset who took the naming vocabulary assessment and scored 33 or below and 3% from the SEED dataset who were not able to take the naming vocabulary assessment due not being able to speak, hear or another reason. The latter figure aligns with Norbury et al (2016) which estimates that approximately 2.34% of children experience "language impairment associated with intellectual disability and/or existing medical diagnosis".

(e.g. medical) to improve their literacy. This intense support is not accounted for in this report as the associations between early years assessments and later outcomes are estimated for an average child, probabilities and costs are likely to be different for this specific group of children, so this specific group is out of the scope of this report.

- PBE's analysis subtract this number of children from the number of children not meeting the literacy standard to isolate just those PBE can reasonably assume can meet the expected level through a more holistic, community-focused approach.
- 3. Calculating the implied impact of literacy on Key Stage 1, and thereby long term economic costs, by local authority considering the number of children that could reasonably reached the expected level in literacy calculated in step 2.
 - PBE has taken the associated monetary value in 2022 prices discounted at age 5 for each later outcome as reported in Table 5 in Annex B for SEN, truancy, and exclusion.
 - For the lifetime earnings value, PBE estimated different costs for foregone earnings to account for regional differences in wages. This process was done by estimating the ratio of each local authority's average wage to the UK average. These ratios were multiplied by the lifetime earnings values in Table 5 to estimate the monetary value of earnings foregone by an average child not reaching the expected level in literacy in each respective local authority. Details of costs by local authority can be found in the spreadsheet Annex D.
 - The sum of the monetary values of SEN, truancy, exclusion and lifetime earnings was then multiplied by the local authority's total points impact at KSI for passing only literacy. This gave us the implied economic cost per child by local authority.
 - The implied economic cost per child by local authority was then multiplied by the number of children who could reasonably reach the expected level in literacy in each Local Authority to estimate the total lifetime cost by area.
- 4. Analysing the implied impact of lower levels of literacy in early years by three levels of deprivation

- PBE used evidence from the Index of Multiple Deprivation on the average deprivation level for each local authority in England.
- Initial analysis suggested that higher levels of deprivation might be correlated with higher levels of children not reaching the literacy standard in EYFS. This prompted PBE to consider modelling this relationship with the deprivation score as a categorical variable, to better understand the nature of this relationship.
- For the regional crosscuts by deprivation, PBE used the nonparametric Kruskal-Wallis test. Wherever the Kruskal-Wallis test showed a statistically significant variation, it was supplemented by the Dunn test which carried out pairwise tests for the three levels of deprivation to identify which groups exhibit statistically significant differences from each other.
- These tests suggested that the following categories of deprivation yielded the strongest, statistically significant differences:
 - i. Low a local authority in the bottom 30% of the distribution of deprivation scores.
 - ii. Medium a local authority in the middle 30% of the distribution of deprivation scores (i.e. between the Low and High categories).
 - iii. High a local authority in the top 40% of the distribution of deprivation scores.
- Comparisons of the number and rate of young children who did not meet the literacy standards, as well as the resulting economic costs, have been presented by these categories.







www.probonoeconomics.com



020 3632 2668